THE DESIGN AND ANALYSIS OF SPLIT-PLOT EXPERIMENTS IN INDUSTRY

By

SCOTT M. KOWALSKI

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS OF THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1999

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Dr. G. Geoffrey Vining for serving as my dissertation advisor. Many thanks go out to him for allowing me the opportunity to serve as an editorial assistant for the *Journal of Quality Technology*. His generosity has given me experiences that most graduate students could only dream of having. Also, not only as he been a mentor for me statistically, but he has also been my friend.

I would also like to thank Dr. John Cornell for his extreme interest in the work that I have done and for being a constant resource for me. In addition, I would like to thank Drs. Jim Hobert, Richard Scheaffer, and Diane Schaub for serving on my committee. Also, I would like to thank Dr. Frank Martin for sitting in on my defense and for developing my interest in Design of Experiments through his course. I extend a big thank you to Dr. Ronald Randles, chairman of the department of statistics, for supporting me through my many years at the University of Florida.

I thank my parents for their endless love and support. They fully believed in me when, at times, I wasn't so sure I would finish. Finally, I have to thank my wife for her love and especially her patience. She has stood by me through everything and I owe her more than I could ever repay.

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements of the Degree of Doctor of Philosophy

THE DESIGN AND ANALYSIS OF SPLIT-PLOT EXPERIMENTS IN INDUSTRY

By

Scott M. Kowalski

December, 1999

Chairman: G. Geoffrey Vining
Major Department: Statistics

Split-plot experiments where the whole plot treatments and the subplot treatments are made up of combinations of two-level factors are considered. Due to cost and/or time constraints, the size of the experiment needs to be kept small. Using fractional factorials and confounding, a method for constructing sixteen run designs is presented. Along with this, semifolding is used to add eight more runs. The resulting twenty-four run design has better estimating properties and gives some degrees of freedom which can be used for estimating the subplot error variance.

Experiments that involve the blending of several components to produce high quality products are known as mixture experiments. In some mixture experiments, the quality of the product depends not only on the relative proportions of the mixture

components but also on the processing conditions. A combined model is proposed which is a compromise between the additive and completely crossed combined mixture by process variable models. Also, a new class of designs that will accomodate the fitting of the new model is considered.

The design and analysis of the mixture experiments with process variables is discussed for both a completely randomized structure and a split-plot structure. When the structure is that of a split-plot experiment, the analysis is more complicated since ordinary least squares is no longer appropriate. With the process variables serving as the whole plot factors, three methods for estimation are compared using a simulation study. These are ordinary least squares (to see how inappropriate it is), restricted maximum likelihood, and using replicate points to get an estimate of pure error. The last method appears to be the best in terms of the increase in the size of the confidence ellipsoid for the parameters and has the added feature of not depending on the model.

CHAPTER 1
INTRODUCTION

A common exercise in the industrial world is that of designing experiments, exploring complex regions, and optimizing processes. The setting usually consists of several input factors that potentially influence some quality characteristic of the process, which is called the response. Box and Wilson (1951) introduced statistical methods to attain optimal settings on the design variables. These methods are commonly known as response surface methodology (RSM), which continues to be an important and active area of research for industrial statisticians.

Many times in industrial experiments, the factors consist of two types: some with levels that are easy to change and one or more with levels that are difficult or costly to change. Suppose for illustration that there is only one factor that is difficult to change. When this is the case, the experimenter usually will fix the level of this factor (ie., restrict the randomization scheme) and then run all combinations or a fraction of all combinations of the other factors, which is known as a split-plot design. Too often, the data obtained from this experiment are analyzed as if the treatment combinations were completely randomized, which can lead to incorrect conclusions as well as a loss of precision. Analysis of data obtained from experiments, such as the example above, need to take the restricted randomization scheme into account.

1.1 Response Surface Methodology

In RSM, the true response of interest, η, can be expressed as a function of one or more controllable factors (at least in the experiment being performed), $\mathbf{x}$, by

$$\eta = g(\mathbf{x}) + \epsilon\,,$$

where the form of the function g is unknown and ϵ is a random error term. The goal is to find, in the smallest number of experiments, the settings among the levels of $\mathbf{x}$ within the region of interest at which η is a maximum or minimum. Because the form of g is unknown, it must be approximated. RSM uses Taylor series expansion to approximate $g(\mathbf{x})$ over some region of interest. Typically, first or second order models are used to approximate $g(\mathbf{x})$. The traditional RSM model would be

$$y_i = f(\mathbf{x})'\boldsymbol{\beta} + \epsilon\,,$$

where

- y_i is the i^{th} response,
- $\mathbf{x}_i$ is the i^{th} setting of the design factors,
- $f(\mathbf{x})$ is the appropriate polynomial expansion of $\mathbf{x}$,
- $\boldsymbol{\beta}$ is a vector of unknown coefficients, and
- the ϵ_i's are assumed to be independent and identically (i.i.d.) distributed as $N(0, \sigma^2)$.

For a more detailed discussion on RSM see Khuri and Cornell (1996), Box and Draper (1987), and Myers and Montgomery (1995).

1.2 Split-Plot Designs

A split-plot design often refers to a design with qualitative factors but can easily handle quantitative factors. Also, a split-plot design usually has replication. However, in the literature it has been common practice to refer to any design that uses one level of restricted randomization regardless of replication as a "split-plot" design. Therefore, in this dissertation, we will use the term split-plot design throughout.

When performing multifactor experiments, there may be situations where complete randomization might not be feasible. A common situation is when the nature of the experiment or factor levels preclude the use of small experimental units. Often a second factor can be studied by dividing the experimental units into sub-units. In these situations, the split-plot experiment can be utilized. The experimental unit is referred to as the whole plot while the sub-units are referred to as the subplots. For every split-plot experiment there are two randomizations. Whole plot treatments are randomly assigned to whole plots based on the whole plot design. Within each whole plot, subplot treatments are randomly assigned to subplots with a separate randomization for each whole plot. This leads to two error terms, one for the whole plot treatments and one for subplot treatments as well as the interaction between whole plot treatments and subplot treatments. Split-plot experiments have been used extensively in agricultural settings. Even so, the following example from Montgomery (1997) shows that there are applications for split-plot experiments in industrial settings.

A paper manufacturer is interested in studying the tensile strength of paper based on three different pulp preparation methods and four cooking temperatures for the

pulp. Each replicate of the full factorial experiment requires 12 observations, and the experimenter will run three replicates. However, the pilot plant is only capable of making 12 runs per day, so the experimenter decides to run one replicate on each of three days. The days are considered blocks. On any day, he conducts the experiment as follows. A batch of pulp is produced by one of the three methods. Then this batch is divided into four samples, and each sample is cooked at one of the four temperatures. Then a second batch of pulp is made using one of the remaining two methods. This second batch is also divided into four samples that are tested at the four temperatures. This is repeated for the remaining method. The data are given in Table 1. This experiment differs from a factorial experiment because of the restriction on the randomization. For the experiment to be considered a factorial experiment, the 12 treatment combinations should be randomly run within each block or day. This is not the case here. In each block a pulp preparation method is randomly chosen, but then all four temperatures are run using this method. For example, suppose method 2 is selected as the first method to be used, then it is impossible for any of the first four runs of the experiment to be, say, method 1, temperature 200. This restriction on the randomization leads to a split-plot experiment with the three pulp preparation methods as the whole plot treatments and the four temperatures as the subplot treatments. It should be noted that conducting a split-plot experiment, as opposed to a completely randomized experiment, can be easier because it reduces the number of times the whole plot treatment is changed. This usually will result in a time savings which will lead to reduced costs. For example, suppose one is interested in six subplot treatments and four whole plot treatments. Let the whole

Table 1: Data for Tensile Strength of Paper (from Montgomery (1997))

	Block 1			Block 2			Block 3		
Pulp Preparation Method	1	2	3	1	2	3	1	2	3
Temperature									
200	30	34	29	28	31	31	31	35	32
225	35	41	26	32	36	30	37	40	34
250	37	38	33	40	42	32	41	39	39
275	36	42	36	41	40	40	40	44	45

plot treatments be comprised of a 2^2 factorial in time and temperature of a kiln. A completely randomized experiment would require the kiln to be fired up quite possibly 24 times. With a split-plot experiment, the kiln only needs to be brought up to the correct temperature 4 times per replicate. This leads to a savings of time and possibly money.

A split-plot experiment can be run inside of many standard designs, such as the completely randomized design (CRD) and the randomized complete block (RCB) design. As in the example from Montgomery (1997), suppose the split-plot experiment is performed using a RCB design. Let Y_{ijk} denote the observation for subplot treatment k receiving whole plot treatment i in block j. Kempthorne (1952) uses as his model

$$y_{ijk} = \mu + \tau_i + \beta_j + \delta_{ij} + \gamma_k + (\tau\gamma)_{ik} + \epsilon_{ijk} \quad \text{for} \quad \begin{array}{l} i = 1, 2, \ldots, t \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, s \end{array}$$

where

- t is the number of levels for the whole plot treatment,
- b is the number of blocks or replicates of the basic whole plot experiment,
- s is the number of levels for the subplot treatment,
- μ is the overall mean,
- τ_i is the effect of the i^{th} whole plot treatment,
- β_j is the effect of the j^{th} block,
- δ_{ij} is the whole plot error term,
- γ_k is the effect of the k^{th} subplot treatment,
- $(\tau\gamma)_{ik}$ is the whole plot treatment by subplot treatment interaction, and
- ϵ_{ijk} is the subplot error.

He uses randomization theory to derive the expected mean squares summarized in Table 2. In this table, σ_δ^2 is the experimental error variance for the whole plot treatments, and σ^2 is the experimental error variance for the subplot treatments.

Many analysts assume that the blocks are random and use an unrestricted mixed model to derive the appropriate mean squares. The most common model for this approach is

$$Y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \gamma_k + (\tau\gamma)_{ik} + \epsilon_{ijk} \qquad (1)$$

$$i = 1, 2, \ldots, t \quad j = 1, 2, \ldots, b \quad k = 1, 2, \ldots, s\,,$$

where

- μ is the overall mean,
- τ_i is the effect of whole plot treatment i,
- β_j is the effect of block j,
- $(\tau\beta)_{ij}$ is the block $\times$ whole plot treatment interaction,
- γ_k is the effect of subplot treatment k,
- $(\tau\gamma)_{ik}$ is the whole plot treatment $\times$ subplot treatment interaction, and
- ϵ_{ijk} is the subplot error.

The $(\tau\beta)_{ij}$ term will be the whole plot error term for the case of an RCB design under the usual assumption of no block $\times$ whole plot treatment interaction. The analysis of variance table associated with the model in Equation (1), assuming whole plot treatments and subplot treatments are fixed and blocks are random, is given in Table 3. If the block by whole plot interaction is called the whole plot error, then Tables 2 and 3 suggest the same basic testing procedures. The following additional constraints and assumptions are needed for hypothesis testing:

$$\sum_i \tau_i = 0, \quad \sum_k \gamma_k = 0\,,$$

$$\beta_j \sim N\left(0, \sigma_\beta^2\right), \quad \delta_{ij} \sim N\left(0, \sigma_\delta^2\right), \quad \epsilon_{ijk} \sim N\left(0, \sigma^2\right),$$

and where δ_{ij} and ϵ_{ijk} are independent.

Montgomery (1997) uses a restricted mixed model as the basis for his analysis of the following form

$$y_{ijkh} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \gamma_k + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + \epsilon_{ijkh},$$

where

- $h = 1, 2, \ldots, r$ is the number of replicates,
- $(\tau\beta)_{ij}$ is the random block by whole plot treatment interaction,
- is the whole plot treatment $\times$ subplot treatment interaction,
- $(\beta\gamma)_{jk}$ is the random block by subplot treatment interaction,
- $(\tau\beta\gamma)_{ijk}$ is the random block by whole plot treatment by subplot treatment interaction.

Under this restricted mixed model, the random interactions involving a fixed factor are assumed subject to the constraint that the sum of that interaction's effects over the levels of the fixed factor is zero. Table 4 gives the resulting expected mean squares, which suggests that there are three distinct error terms. The block by whole plot by subplot interaction is used to test the whole plot by subplot interaction; the block by subplot interaction is used to test the main effect of the subplot treatment; and the block by whole plot interaction is used to test the main effect of the whole plot.

Table 2: Expected Mean Squares Table Under Randomization Theory

Source	df	Expected Mean Square
Whole Plot Treatment	$t-1$	$\sigma_\delta^2 + \frac{bs}{t-1}\sum_{i=1}^{t}\tau_i^2$
Blocks	$b-1$	$\sigma^2 + st\sigma_\beta^2$
Whole Plot Error	$(t-1)(b-1)$	σ_δ^2
Subplot Treatment	$s-1$	$\sigma^2 + \frac{bt}{s-1}\sum_{k=1}^{s}\gamma_k^2$
Whole Plot × Subplot	$(t-1)(s-1)$	$\sigma^2 + \frac{b}{(t-1)(s-1)}\sum_{i=1}^{t}\sum_{k=1}^{s}(\tau\gamma)_{ik}^2$
Subplot Error	$t(b-1)(s-1)$	σ^2

Note, if $h = 1$, the variance of ϵ_{ijkh} is not estimable. This restricted analysis reduces to the other two analyses only if the block by subplot interaction is unimportant. In such a case, its contribution can be pooled with the block by whole plot by subplot interaction to form the same error term as the randomization and unrestricted mixed model analyses.

Whole plot treatments are applied to blocks of t units which can be divided further into s subunits, where s is the number of levels of the subplot treatment. Any differences among these blocks must be confounded with the whole plot treatment comparisons. Consequently, comparisons among the subplot treatments are made with greater precision, and this leads to the more important factor usually being assigned to the subplot. Using the unrestricted model and Table 3, it is seen that the null hypothesis of no whole plot treatment effect, $H_0 : \tau_1 = \tau_2 = \cdots = \tau_t$, versus at least one not equal, is tested using the Block × Whole Plot Treatment interaction as

Table 3: Expected Mean Squares Table Under the Most Common Unrestricted Mixed Model

Source	df	Expected Mean Square
Whole Plot Treatment	$t-1$	$\sigma^2 + s\sigma^2_\delta + \frac{bs}{t-1}\sum_{i=1}^{t}\tau_i^2$
Blocks	$b-1$	$\sigma^2 + s\sigma^2_\delta + st\sigma^2_\beta$
Block × Whole Plot Treatment	$(t-1)(b-1)$	$\sigma^2 + s\sigma^2_\delta$
Subplot Treatment	$s-1$	$\sigma^2 + \frac{bt}{s-1}\sum_{k=1}^{s}\gamma_k^2$
Whole Plot × Subplot	$(t-1)(s-1)$	$\sigma^2 + \frac{b}{(t-1)(s-1)}\sum_{i=1}^{t}\sum_{k=1}^{s}(\tau\gamma)_{ik}^2$
Error	$t(b-1)(s-1)$	σ^2

Note: Whole Plot and Subplot Treatments are assumed fixed while Blocks are assumed to be random.

Table 4: Expected Mean Squares Table Under the Restricted Mixed Model

Source	df	Expected Mean Square
Whole Plot Treatment	$t-1$	$\sigma^2 + s\sigma^2_{\tau\beta} + \frac{bs}{t-1}\sum_{i=1}^{t}\tau_i^2$
Blocks	$b-1$	$\sigma^2 + s\sigma^2_{\tau\beta} + st\sigma^2_\beta$
Block × Whole	$(t-1)(b-1)$	$\sigma^2 + s\sigma^2_{\tau\beta}$
Subplot Treatment	$s-1$	$\sigma^2 + t\sigma^2_{\beta\gamma} + \frac{bt}{s-1}\sum_{k=1}^{s}\gamma_k^2$
Block × Sub	$(b-1)(s-1)$	$\sigma^2 + t\sigma^2_{\beta\gamma}$
Whole Plot × Subplot	$(t-1)(s-1)$	$\sigma^2 + \sigma^2_{\tau\beta\gamma} + \frac{b}{(t-1)(s-1)}\sum_{i=1}^{t}\sum_{k=1}^{s}(\tau\gamma)_{ik}^2$
Block × Whole × Sub	$(t-1)(b-1)(s-1)$	$\sigma^2 + \sigma^2_{\tau\beta\gamma}$

the error term. The hypothesis of no subplot treatment effect, $H_0 : \gamma_1 = \gamma_2 = \cdots = \gamma_s$, versus at least one not equal, is tested using Error which is also used to test the significance of the whole plot $\times$ subplot treatment interaction.

Suppose the whole plot and subplot treatments are a factorial structure. In this case, after the hypothesis tests above are performed, a more detailed investigation of the individual factors and their interactions can be carried out. For example, consider the situation discussed above with $t = 4$ whole plot treatments consisting of a 2^2 factorial in z_1 and z_2 and $s = 4$ subplot treatments also consisting of a 2^2 factorial in x_1 and x_2. The $t - 1 = 3$ degrees of freedom (df) for the whole plot treatments can be partitioned into single df contrasts z_1, z_2, and $z_1 z_2$. Likewise, the $s - 1 = 3$ df for the subplot treatments can be partitioned into a single df contrasts x_1, x_2, and $x_1 x_2$. Also, the $(t - 1)(s - 1) = 9$ df for the whole plot $\times$ subplot treatment interaction can be broken down into 9 single df effects involving z_1, z_2, x_1, and x_2 (see Table 5). Orthogonal contrasts should be calculated and tested for each factor and the interactions using the appropriate error term from the original analysis. This can be accomplished in SAS by using PROC GLM and the CONTRAST statement along with the option E = *error term* after the model statement for the whole factors and interactions. For the subplot factors, interactions among subplot factors, and whole plot $\times$ subplot factor interactions, the analysis can be run a second time. In this second analysis, the treatments in the model statement can be entered as factors and interactions, similar to a regression model. The correct tests for the subplot factors and whole plot $\times$ subplot factor interactions are given by SAS.

Table 5: Analysis of Variance Table for a Split-Plot Experiment Run Using a RCB Design With Factorial Structure and the Most Common Unrestricted Model

Source	df
Whole Plot Treatment	$t-1=3$
z_1 †	1
z_2 †	1
$z_1 z_2$ †	1
Blocks	$b-1$
Block × Whole Plot Treatment	$(t-1)(b-1)$
Subplot Treatment	$s-1=3$
x_1 ††	1
x_2 ††	1
$x_1 x_2$ ††	1
Whole Plot × Subplot	$(t-1)(s-1)=9$
$z_1 x_1$ ††	1
$z_1 x_2$ ††	1
$z_2 x_1$ ††	1
$z_2 x_2$ ††	1
$z_1 x_1 x_2$ ††	1
$z_2 x_1 x_2$ ††	1
$z_1 z_2 x_1$ ††	1
$z_1 z_2 x_2$ ††	1
$z_1 z_2 x_1 x_2$ ††	1
Error	$t(b-1)(s-1)$

† These terms are tested using the Block × Whole Plot Treatment interaction.
†† These terms are tested using Error.

The concept of the split-plot design can be extended if further randomization restrictions exist. For example, suppose there are two levels of randomization restrictions within a block in which case we might have a split-split-plot design. For a more detailed discussion of split-plot designs and their extensions see Yates (1937), Cox (1958), Wooding (1973) and Montgomery (1997).

1.3 Dissertation Goals

The focus of this dissertation is to enhance our understanding of the design and analysis of split-plot experiments. The experiments considered will be industrial in nature. As much as possible, the dissertation will focus on or discuss the types of experiments that would be run in industry in terms of size and resources. An important goal is to come up with methods that are clear, practical, and easy to implement. In other words, this dissertation will address issues of real concern to applied statisticians working in industry and provide them some tools that can be used with split-plot experiments. Below two industrial statisticians have been kind enough to share real situations that help to show the relevance of the work in this dissertation.

A Food Industry Example

Frozen heat-and-serve pastries, along with shelf-stable ready-to-eat pastries, represent a large segment of the convenience foods that today's consumers crave. Optimized proofing and baking operations are critical to the successful manufacture of high quality baked goods such as these. However, as this market segment has grown,

so has the manufacturing capacity, which has necessitated the installation of new proofers and ovens. Given the complexity of these operations, qualifying a new piece of proofing or baking equipment poses a challenging experimental design problem: how do you design an experiment to explore the operating profile for a new proofer or oven?

As an example, consider a continuous oven, in which dough-based products move through on a belt. The oven has two zones, which are controlled independently. In each zone you can adjust the Temperature, the Relative Humidity (RH), the Air Flow Speed (AF), and the Residence Time. In general, the conditions in each zone will be different, as each zone is used to impart different characteristics to the product. All of these variables will impact the quality of the finished product.

Experimenting with this type of oven requires a restricted randomization. You can easily reset the air flow and residence time in each zone on the fly, but changes in the temperature and relative humidity require a waiting period to allow the oven to return to steady state. Thus, oven experiments are typically conducted as a split-plot design with four whole plot treatments, namely

Zone 1 Temp, Zone 1 RH, Zone 2 Temp, and Zone 2 RH,

and four split plot treatments, namely

Zone 1 AF, Zone 1 Res Time, Zone 2 AF, and Zone 2 Res Time.

In addition, we typically want to evaluate the effect of the oven on at least two products (say Products A and B) to see if they respond differently to the oven. Each

product that is baked is evaluated in a variety of ways, including sensory characterizations and analytical and physical testing.

What makes this experimental setup so difficult, is that exploring the profile of a new oven must typically occur on prototype equipment at the oven manufacturer. This means the experiment must be conducted in a very short period of time, often two days or less. This makes it imperative that the experiment have as few runs as possible usually between 12 and 20 runs. The research described in this thesis is directly applicable to problems like this, and will be very useful for teams of process engineers charged with gathering the data they need to fully evaluate candidate ovens and proofers.

An Integrated Circuits Example

Integrated microflex circuits are manufactured over several, very complicated process steps. Circuit plating is a key step in this process. It involves depositing a uniform layer of copper on the microflex circuits. Copper thickness is a key quality characteristic due to functionality issues. High variability in copper thickness results in poor bonding of chips to these circuits. Some of the variables that effects circuit thickness are the circuit geometry, the line speed, the current in amperes, the copper concentration in the chemistry bath, and the concentration of both sulfuric acid and hydrogen peroxide.

Designing experiments to optimize copper thickness is a challenge because of the presence of hard-to-change variables. In particular, restricted randomization occurs with circuit geometry, line speed, and current. Randomization is not restricted for

the remaining variables. Data from such experiments are analyzed by assuming that it arose from a completely randomized experiment. Research in the area of split-plot experiments with multiple whole plot and subplot factors is lacking and the work in this dissertation should be of real help to industrial statisticians.

1.4 Overview

The literature review that follows in Chapter 2 is intended to familiarize the reader with other work that discusses split-plot designs in RSM. Chapter 3 begins with some background information on 2^k factorial experiments and the remainder of Chapter 3 is devoted to a more in-depth look at confounding in split-plot experiments. In Chapter 4, a new model and class of designs for mixture experiments with process variables will be developed in a completely randomized setting. Finally, the last chapter will assume a split-plot structure for the mixture experiments with process variables described in Chapter 4.

CHAPTER 2
LITERATURE REVIEW

The split-plot error structure has been underutilized in RSM. Most RSM experiments assume a completely randomized error structure. Letsinger, Myers, and Lentner (1996, pg. 382) point out, "Unfortunately, while this completely randomized assumption simplifies analysis and research, independent resetting of variable levels for each design run may not be feasible due not only to equipment and resource constraints, but also budget restrictions." This chapter focuses on the literature involving restricted randomization within RSM.

2.1 Split-Plot Confounding

When the whole plot and/or subplot treatments are of a factorial nature, it is possible to reduce the number of whole plots and/or subplots needed through fractionating. This is important in industrial experiments where constraints limit the size of the experiment. Bartlett (1935) suggested the possibility of confounding higher-order subplot interactions to reduce the number of subplots needed within each whole plot. Later, split-plot confounding was studied by Addelman (1964). He provided a table containing factorial and fractional-factorial arrangements that involve split-plot confounding. However, he did not consider confounding within the whole plots. Letsinger, Myers, and Lentner (1996) discuss the possibility of split-plot confounding

with the use of their noncrossed bi-randomization designs. Box and Jones (1992) illustrate split-plot confounding using a cake mix example.

In some experiments, there are constraints on the number of subplots within each whole plot. When the whole plots are arranged in a CRD, Robinson (1967) discussed situations where the number of subplots per whole plot is less than the number of subplot treatments. The whole plots are treated as blocks and then a balanced incomplete block (BIB) design is used to allocate the subplot treatments to the whole plots. If the whole plots are arranged in an RCB design, the same procedure can be applied. If the number of whole plots per block is less than the number of whole plot treatments, then an incomplete block design can be used there as well. Robinson (1970) gave details on the case when both whole plot and subplot treatments are arranged in incomplete block designs. Essentially, the procedure amounts to arranging the whole plot treatments in blocks using a BIB design and then considering the whole plots as blocks and arranging the subplot treatments in another BIB design. Robinson (1970) provided formulas for the estimates of the main effects and interactions for three cases: within whole plot, between whole plot within block, and between blocks. Formulas are also given for the variance of the differences of these estimates for each case.

Huang, Chen and Voelkel (1998) also investigate fractionating two-level split-plot designs at both the whole plots and the subplots. They consider $2^{(n_1+n_2)-(k_1+k_2)}$ split-plot designs which are associated with a subset of the 2^{n-k} fractional factorial designs where $n = n_1 + n_2$ and $k = k_1 + k_2$. The criterion used to select the optimal design is that of minimum-aberration which is the design that has smallest number of

words in the defining contrast with the fewest letters. Two methods are presented for constructing minimum-aberration split-plot designs. The first method decomposes the 2^{n-k} design into the $2^{(n_1+n_2)-(k_1+k_2)}$ split-plot design. This method is used to derive extensive, though incomplete, tables of the designs. The second and more complicated method which involves linear integer programming is used when the first method fails.

Minimum-aberration two-level split-plot designs are also discussed in Bingham and Sitter (1999). A combined search and sequential algorithm is presented for constructing all non-isomorphic minimum-aberration split-plot designs which include the tables of Huang, Chen and Voelkel (1998). Bingham and Sitter (1999) catalog designs for 16 and 32 runs containing up to 10 factors. Included in this catalog are the second and third best minimum-aberration designs since sometimes it may be desirable to use these designs.

2.2 Split-Plots in Robust Parameter Designs

Genichi Taguchi proposed methods for designing experiments for product design that are robust to environmental variables. The goal of robust design is to design an experiment that identifies the settings of the design factors that make the product robust to the effects of the noise variables. The design factors, which are factors controlled during manufacturing, make up the inner array while the environmental factors, or noise factors, make up the outer array. Environmental factors are factors that are difficult to control and can cause variation in the use or performance of products. The experimental design or "crossed array" consists of crossing each

experimental design setting of the inner array with each experimental design setting of the outer array. Unless the number of factors in these arrays is small, Taguchi's designs become large and expensive.

An alternative to Taguchi's crossed array is the "combined" array. The combined array utilizes a single experimental design in both the design and environmental factors. Therefore, the response is modeled directly as a function of the design factors and the environmental factors using a single linear model. More details on the combined array can be found in Welch, Kang, and Sacks (1990); Shoemaker, Tsui, and Wu (1991); and O'Donnell and Vining (1997).

Bisgaard (1999) discusses split-plot designs in association with inner and outer array designs. He focuses on screening experiments that use restricted randomization. The paper gives a nice overview of defining relations and confounding structures for the $2^{k-p} \times 2^{q-r}$ split-plot designs. In addition to split-plot confounding, Bisgaard (1999) points out that the same fraction of the subplot factors can be run in each whole plot. The appropriate standard errors for testing effects when using split-plot confounding are also given.

Box and Jones (1992) investigate the use of split-plot designs as an alternative to the crossed array. They consider three experimental arrangements where the robust parameter design is set up as a split-plot design:

1. arrangement (a) — the whole plots contain the environmental factors and the subplots contain the design factors;

2. arrangement (b) — the whole plots contain the design factors and the subplots contain the environmental factors;

3. arrangement (c) — the subplot factors are assigned in "strips" across the whole plot factors (commonly called a strip-block experiment).

These three arrangements are illustrated through an example seeking the best recipe for a cake mix. Three design factors have been identified as affecting taste. They are flour, shortening, and egg powder and are studied using a 2^3 factorial design. The consumer may have an oven in which the temperature is biased up or down. Also, the consumer may overcook or undercook the cake. Therefore, the recipe is to be robust to two environmental factors, oven temperature and baking time, whose levels are combined using a 2^2 factorial design.

Arrangement (a)

Under this arrangement, the whole plots contain the environmental factors and the subplots contain the design factors. Suppose there are m levels of the environmental factors, $E_1, E_2, \ldots, E_j, \ldots, E_m$, applied to the whole plots, n levels of the design factors, $D_1, D_2, \ldots, D_i, \ldots, D_n$, applied to the subplots, and l replicates, $r_1, r_2, \ldots, r_u, \ldots, r_l$, with the whole plots in l randomized blocks. For the cake mix example, $m = 4$, $n = 8$, and $l = 1$. Arrangement (a) requires $m \times n \times l$ subplots and $m \times l$ whole plots. Thus for the cake mix example, $4 \times 8 \times 1 = 32$ cake mix batches are required, but only $4 \times 1 = 4$ operations of the oven are necessary. By comparison, a completely randomized cross-product array would require 32 cake mix

batches and 32 operations of the oven. Thus, the split-plot arrangement has saved time by reducing the number of operations of the oven.

The model for arrangement (a) is

$$y_{ijk} = \mu + \gamma_k + \alpha_j + \eta_{jk} + \delta_i + (\alpha\delta)_{ij} + \epsilon_{ijk}\,,$$

where y_{ijk} is the response of the k^{th} replicate of the i^{th} level of factor D and the j^{th} level of factor E, μ is the overall mean, γ_k is the random effect of the k^{th} replicate with $\gamma_k \sim N(0, \sigma_\gamma^2)$, α_j is the fixed effect of the j^{th} level of factor E, δ_i is the fixed effect of the i^{th} level of factor D, $(\alpha\delta)_{ij}$ is the interaction effect of the i^{th} level of D and the j^{th} level of E, $\eta_{jk} \sim N(0, \sigma_w^2)$ is the whole plot error, $\epsilon_{ijk} \sim N(0, \sigma_s^2)$ is the subplot error, and η_{jk} and ϵ_{ijk} are independent.

Arrangement (b)

With this arrangement, the whole plots contain the design factors while the subplots contain the environmental factors. Arrangement (b) requires only $8 \times 1 = 8$ cake mix batches but requires $4 \times 8 \times 1 = 32$ operations of the oven. Again, a completely randomized cross-product array would use 32 cake mix batches and 32 operations of the oven. Here, the savings of the split-plot design is not as great since only the number of cake mix batches is reduced. This is not an ideal situation for industrial experiments. First of all, the design factors are of greater interest. Therefore, applying the design factors to the whole plots results in a loss of precision for the design factors. Hence, it is possible to have large differences between the levels of the design factors that are insignificant when tested. Also, from an economic point

of view, arrangement (b) is costly. It requires an inefficient use of the environmental factors which in industrial experiments are typically the difficult or expensive to change factors.

The model for arrangement (b) is

$$y_{ijk} = \mu + \gamma_k + \delta_i + \eta_{ik} + \alpha_j + (\alpha\delta)_{ij} + \epsilon_{ijk} ,$$

where y_{ijk} is the response of the k^{th} replicate of the i^{th} level of factor D and the j^{th} level of factor E, μ is the overall mean, γ_k is the random effect of the k^{th} replicate with $\gamma_k \sim N(0, \sigma_\gamma^2)$, α_j is the fixed effect of the j^{th} level of factor E, δ_i is the fixed effect of the i^{th} level of factor D, $(\alpha\delta)_{ij}$ is the interaction effect of the i^{th} level of D and the j^{th} level of E, $\theta_{ik} \sim N(0, \sigma_w^2)$ is the whole plot error, $\epsilon_{ijk} \sim N(0, \sigma_s^2)$ is the subplot error, and θ_{ik} and ϵ_{ijk} are independent.

Arrangement (c)

Now, consider the arrangement where the subplot treatments are randomly assigned in strips across each block of whole plot treatments (see Table 6). For the cake mix example, suppose each of the $n = 8$ batches of cake mix is subdivided into $m = 4$ subgroups. One subgroup from each batch is then selected, and these eight are baked in the same oven at the appropriate temperature for the appropriate time. This arrangement requires only 8 cake mix batches and only 4 operations of the oven. Therefore, the strip-block experiment is easier to run than the completely randomized cross-product design, as well as both arrangements (a) and (b).

Table 6: Strip-Block Arrangement (Box and Jones (1992))

	Block 1				Block 2				Block 3			
	a_1	a_2	a_3		a_3	a_2	a_1		a_1	a_3	a_2	
b_1				b_2				b_1				
b_2				b_1				b_2				

The model for the strip-block arrangement is

$$y_{ijk} = \mu + \gamma_k + \alpha_j + \eta_{jk} + \delta_i + \theta_{ik} + (\alpha\delta)_{ij} + \epsilon_{ijk}\,,$$

where y_{ijk} is the response of the k^{th} replicate of the i^{th} level of factor D and the j^{th} level of factor E, μ is the overall mean, γ_k is the random effect of the k^{th} replicate with $\gamma_k \sim N(0, \sigma_\gamma^2)$, α_j is the fixed effect of the j^{th} level of factor E, δ_i is the fixed effect of the i^{th} level of factor D, $(\alpha\delta)_{ij}$ is the interaction effect of the i^{th} level of D and the j^{th} level of E. In arrangement (c), $\eta_{jk} \sim N(0, \sigma_E^2)$, $\theta_{ik} \sim N(0, \sigma_D^2)$, $\epsilon_{ijk} \sim N(0, \sigma^2)$ and η_{jk}, θ_{ik} and ϵ_{ijk} are independent.

ANOVA tables for all three arrangements are given in Box and Jones (1992). These tables indicate the appropriate denominators for tests involving the design factors, the environmental factors, and their interactions assuming replication. When there is no replication, normal probability plots, one for the whole plot factors and one for the subplot factors and whole plot $\times$ subplot interactions, can be used to select significant effects. Also, if the design and environmental factors are factorial combinations, it may be possible to pool negligible higher order interactions to get estimates of the whole plot and subplot errors.

It is of great interest to the researcher to learn how and which environmental factors influence the design variables. This information is contained in the subplot $\times$ whole plot interactions. However, Taguchi's analysis is commonly conducted in terms of a performance statistic, such as the signal to noise ratio (SNR). The SNR is calculated for each point in the inner array using data obtained from the outer array about that point. Therefore, Taguchi ignores any information contained in the interactions of the design and environmental factors. This is generally considered to be a serious drawback to the Taguchi analysis.

Phadke (1989) presented an example involving a polysilicon deposition process which he analyzed using Taguchi's SNR's. Polysilicon film is typically deposited on top of the oxide layer of the wafers using a hot-wall, reduced pressure reactor. The reactant gases are introduced into one end of a three-zone furnace tube and are pumped into the other end. The wafers enter the low-pressure chemical vapor deposition furnace in two quartz boats, each with 25 wafers, and polysilicon is deposited simultaneously on all 50 wafers. The desired output of this process is a wafer which has a uniform layer of film of a specified thickness. Six design factors each at three levels were identified: temperature, pressure, nitrogen flow, silane flow, setting time, and cleaning method. Tube location and die location were considered noise factors. Three responses, film thickness, particle counts, and deposition rate, were of interest. The smaller the better SNR was used in the analysis for particles, the target is best SNR was used for thickness, and a 20 $\log_{10}$ transformation was used for deposition rate. The data were analyzed using ANOVA techniques to determine the effect of

each design factor on the responses. A more detailed discussion of the selection of factors, design, and analysis of the SNR's is contained in Chapter 4 of Phadke (1989).

The actual structure of this experiment was a split-split-plot design because there are three sizes of experimental units with different sources of variation. The design factors are applied to the tube (run-to-run variability); the location in the tube affects the wafer (wafer-to-wafer variability), whereas location in the wafer affects the die (die-to-die variability). Therefore, using Taguchi's SNR's to analyze this experiment will result in a complete loss of information in the design $\times$ noise factor interactions. Cantell and Ramírez (1994) reanalyzed the data as if it were a split-split-plot design. They pooled higher order interactions to get the necessary error terms in order to perform hypothesis tests on the design factors and the design $\times$ noise factor interactions. Interaction plots were used to determine the level of the design factor that minimized the variation across the levels of the noise factors. Although the final recommendations on the design factor levels by Cantell and Ramírez (1994) differed from Phadke (1989) on only one of the six design factors, the use of the split-split-plot design has allowed the process engineer to have a better understanding of the sources of variation. This added information may lead to process improvement in the future.

Kempthorne (1952) and Box and Jones (1992) provide details on the relative efficiency of these split-plot designs compared to the CRD and RCB. A summary of their conclusions is provided here. Consider the split-plot experiment as a uniformity trial. If the uniformity trial was run as a CRD or a RCB experiment, then, for arrangements (a) and (b), the subplot factor effects and the subplot $\times$ whole plot interactions are estimated more precisely than the whole plot factor effects. Compared

with arrangements (a) and (b), the strip-block design estimates the subplot × whole plot interactions more precisely but the subplot factor effects with less precision. However, the whole plot factor effects are estimated with equal precision. Based on these results, arrangement (a) with the environmental factors applied to the whole plots is generally preferred over arrangement (b). Both the strip-block design and the split-plot design with the design factors applied to the subplots can be extremely useful in robust parameter design.

2.3 Bi-Randomization Designs

Letsinger, Myers, and Lentner (1996) introduced bi-randomization designs (BRD's). BRD's refer to designs with two randomizations similar to that of a split-plot design. The whole plot variables are denoted by $\mathbf{z} = (z_1, z_2, \ldots, z_z)$ while the sub-plot variables are denoted by $\mathbf{x} = (x_1, x_2, \ldots, x_x)$. Hence, the i^{th} design run is $(\mathbf{z}_i, \mathbf{x}_i)$. BRD's are broken into two classes, crossed and non-crossed. Crossed BRD's are constructed as follows:

1. randomize the a unique combinations of $\mathbf{z}$ to the whole plot experimental units (EU's), then

2. randomize the b levels of $\mathbf{x}$ to the smaller EU's within each whole plot (see Table 7).

Thus every level of $\mathbf{x}$ is "crossed" with every level of $\mathbf{z}$. These designs are the usual split-plot designs.

Table 7: Crossed BRD From Letsinger, Myers, and Lentner (1996)

$\mathbf{z}_1$	$\mathbf{x}_1$	$\cdots$	$\mathbf{x}_b$
$\mathbf{z}_2$	$\mathbf{x}_1$	$\cdots$	$\mathbf{x}_b$
$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\mathbf{z}_a$	$\mathbf{x}_1$	$\cdots$	$\mathbf{x}_b$

Table 8: Noncrossed BRD From Letsinger, Myers, and Lentner (1996)

$\mathbf{z}_1$	$\mathbf{x}_{11}$	$\cdots$	$\mathbf{x}_{1b_1}$
$\mathbf{z}_2$	$\mathbf{x}_{21}$	$\cdots$	$\mathbf{x}_{2b_2}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\mathbf{z}_a$	$\mathbf{x}_{a1}$	$\cdots$	$\mathbf{x}_{ab_a}$

The non-crossed BRD's differ from the crossed BRD's in that not all levels of $\mathbf{x}$ are associated with $\mathbf{z}_i$. The whole plots have different levels of the sub-plots and need not have the same number of levels. Non-crossed BRD's are constructed as follows:

1. randomize the a unique combinations of $\mathbf{z}$ to the whole plot EU's, then
2. randomize the b_i levels of $\mathbf{x}$ to the smaller EU's within each whole plot (see Table 8).

The distinction between these two can be thought of in terms of the sub-plot factors. The crossed BRD might be represented by a 2^k factorial in the sub-plot factors while the non-crossed BRD might use a 2^{k-p} fractional-factorial in the sub-plot factors but not the same 2^{k-p} set of treatments.

For both crossed and non-crossed BRD's, the two randomizations complicate the error structure. The first randomization leads to the whole plot error variance, σ_δ^2, while the second randomization leads to the sub-plot variance, σ_ϵ^2. It is assumed that the covariance between any two observations on the same whole plot is constant over all whole plots and that observations on two sub-plots from different whole plots are uncorrelated. The response surface model is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\delta} + \boldsymbol{\epsilon}\,,$$

where $\boldsymbol{\delta} + \boldsymbol{\epsilon} \sim N(0, \mathbf{V})$ with $\mathbf{V} = \sigma_\delta^2\mathbf{J} + \sigma_\epsilon^2\mathbf{I}$, where $\mathbf{J}$ is a block diagonal matrix of $\mathbf{1}_{b_i} \times \mathbf{1}'_{b_i}$, and where b_i is the number of observations within the i^{th} whole plot. Now using generalized least squares (GLS), the maximum likelihood estimate (MLE) of the response surface model is

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \tag{2}$$

with

$$\text{Var}\left(\hat{\boldsymbol{\beta}}\right) = \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}. \tag{3}$$

From Equation (2), it is seen that the model estimation depends on the matrix $\mathbf{V}$ and thus both σ_δ^2 and σ_ϵ^2.

Suppose that the response surface model is partitioned into the whole plot and sub-plot terms as

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\gamma} + \mathbf{X}^*\boldsymbol{\beta}^* + \mathbf{Z}'\boldsymbol{\Delta}\mathbf{X}^*\,,$$

where $\boldsymbol{\Delta}$ is a matrix of whole plot $\times$ sub-plot interaction parameters. The response surface design should be large enough to test for general lack of fit as well as lack of

fit from the whole plots. Therefore, the number of whole plots available must exceed the number of parameters in $\boldsymbol{\gamma}$.

For the crossed BRD, there is an equivalence between ordinary least squares (OLS) and GLS. This equivalence means that Equation (2) becomes

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

and the model estimation no longer depends on the error variance. However, for testing purposes, the error variance must be estimated. Letsinger, Myers, and Lentner (1996) suggest augmenting the response surface model with insignificant whole plot terms, $\mathbf{Z}^*\boldsymbol{\rho}$, to saturate the $a-1$ whole plot degrees of freedom. The whole plot saturated model can be used to calculate lack of fit sums of squares for both the whole plots and the sub-plots. Then approximate t-tests can be formed by substituting the estimated error variances into Equation (3).

Non-crossed BRD's present a more complicated situation. The equivalency of OLS and GLS is only true in the case of a first-order model. Although more complex, the above method can be adapted for the first-order case. Letsinger, Myers, and Lentner (1996) compare three methods for the second-order case. They are OLS, iterative reweighted least squares (IRLS), and restricted maximum likelihood (REML). Though IRLS and REML appear to be better methods, the "best" method depends on the design, model, and any prior information.

Bi-randomization introduces the need for new definitions for design efficiency because efficient designs in the literature are based on a completely randomized error structure. For example, for the BRD the D-optimality criterion (see, eg., Kiefer and

Wolfowitz (1959)) becomes

$$\min_{D} \left| N \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} \right)^{-1} \right|$$

over all designs D. Letsinger, Myers, and Lentner (1996) provide comparisons of several first- and second-order designs. For the second-order designs, the popular central composite design (CCD) proves to be a good design.

2.4 Split-Plots in Industrial Experiments

Lucas and Ju (1992) investigated the use of split-plot designs in industrial experiments where one factor was difficult to change and its levels served as the whole plot treatments. They began their study with a simulation exercise using a four factor face-centered cube design with four center points. They let x_1 correspond to the hard-to-change factor, while x_2, x_3, and x_4 were easy to vary. This design allowed for the fitting of the quadratic model

$$Y = \beta_0 + \sum_{i=1}^{4} \beta_i x_i + \sum_{i=1}^{4} \beta_{ii} x_i^2 + \sum_{i=1}^{3} \sum_{j=i+1}^{4} \beta_{ij} x_i x_j + \epsilon \,.$$

However, since the error was the only term of interest, all the regression coefficients can be zero. Therefore, data was generated using

$$Y = \epsilon = \epsilon_w + \epsilon_S \,,$$

where $\epsilon_w \sim N(0, \sigma_w^2)$ was the error term associated with changing the level of x_1 and $\epsilon_s \sim N(0, \sigma_s^2)$ was the error associated with any new experimental run. Twentyeight runs were generated using the following steps:

1. Generate $\epsilon_w \sim N(0, \sigma_w^2)$ and $\epsilon_s \sim N(0, \sigma_s^2)$.

2. $Y = \epsilon_w + \epsilon_s$.

3. If the level of x_1 of the current run is different from that of the previous run, a new value of both ϵ_w and ϵ_s is generated. Otherwise, generate a new value for ϵ_s only.

4. Go to step 2 until all 28 runs are completed.

The data were generated for completely randomized, completely restricted, and partially restricted run orders. For a partially restricted run order, each level of the hard-to-change factor was visited exactly twice and the runs at each level were randomly divided into two equal groups. Each time a data set was generated, the least squares estimates of the β's were computed and the residual error was estimated. The simulation procedure was repeated 1,000 times.

Lucas and Ju (1992) summarized their simulation results in a table with a listing of the standard deviations of the regression coefficients for the three different ways of running the experiment. The restricted randomization case has a much smaller residual standard deviation and much smaller standard deviations for all the regression coefficients except those associated with the hard-to-change factor, β_1 and β_{11}. These results correspond with the general result that split-plot designs will produce increased precision on the subplot factors while sacrificing precision on the whole plot factors. The magnitudes of the coefficients of the estimated standard deviations for the partially restricted case were greater than those with the completely randomized case but less than the corresponding estimates for the completely restricted case.

A similar simulation was conducted for two-level factorials (see Lucas and Ju (1992)). They considered a 2^4 factorial with x_1 as the hard-to-change factor. This allows the fitting of a regression model that includes the linear and interaction terms. Again, a summary table is provided by Lucas and Ju (1992) showing s imilar results to the other experimental scenarios. The completely restricted experiment had smaller standard deviations for all the regression coefficients except β_1. Table 9 gives the formula for the variance of the regression coefficients for a 2^k factorial experiment with one hard-to-change factor. Recall that in the partial restricted case, the blocking was done at random. This can be improved on by blocking orthogonally. The 2^4 factorial can easily be blocked orthogonally in 4 blocks of size 4 or 8 blocks of size 2. Both of these blocking schemes are an improvement over the partially restricted case in that they have smaller standard deviations on the easy-to-vary factors.

Cornell (1988) discusses the analysis of data from mixture experiments with process variables where the mixture blends are embedded in the process variable combinations as in "a split-plot design". The mixture process variables are factors that are not mixture ingredients but whose levels could affect the blending properties of the mixture components. To illustrate this situation, Cornell uses an example from Cornell and Gorman (1984) involving fish patties. The mixture experiment involves making fish patties from different blends of three fish species (mullet, sheepshead, and croaker). The patties were subjected to factor level combinations of three process variables (cooking temperature, cooking time, and deep-frying time). Each process variable was studied at two levels. When process variables are included in a mixture

experiment, complete randomization tends to be impractical. This leads to a restriction on randomization and lends itself to the split-plot design.

Cornell (1988) considers factor-level combinations of the process variables as the whole plot treatments and the mixture component blends as the subplot treatments, but points out that their roles can be switched. Hence, a combination of the levels of the process variables is selected and all blends are run at this combination. Another combination of the process variable levels is chosen and all blends are run at this combination. This procedure is continued until all combinations of the process variables are performed. Following a replication of the complete design, the split-plot nature of the experiment leads to two error terms which are used to assess the significance of the effects of the whole plot treatments, the subplot treatments, and their interaction. Several regression-type models are considered for estimating the effects of the process variables, the blending properties of the mixture components, and interactions between the two. The paper explains how to estimate the regression coefficients as well as how to obtain variances and perform hypotheses tests. Both balanced and unbalanced cases are considered. The hypothesis testing procedures are illustrated with two completely worked-out numerical examples.

Santer and Pan (1997) discuss subset selection procedures for screening in two-factor treatment designs. The paper deals mainly with split-plot designs run in complete blocks; however, the strip-plot design is also discussed. One factor serves as the whole plot factor while the other is the subplot factor. The goal is to select a subset of the treatment combinations associated with the largest mean. Subset selection procedures are given for additive and nonadditive factor cases, where neither of the

Table 9: Variance of the Regression Coefficients For a 2^k With One Hard-To-Change Factor (from Lucas and Ju (1992))

$$\mathrm{Var}(b) = \mathrm{A}\sigma_w^2 + \mathrm{B}\sigma_s^2$$

Hard To Change Variable		Other Terms	
A	B	A	B
$\frac{1-P}{2^k} + \frac{P}{2}$	$\frac{1}{2^k}$	$\frac{1-P}{2^k}$	$\frac{1}{2^k}$

$P = 1/(2^{k-2}+1)$ for the completely randomized design.
$P = 1$ for the completely restricted design.
$P = (2^{k-1}-2)/[2(2^{k-1}-1)]$ for the partially restricted design.

procedures depend on the block variance.

Miller (1997) considers various fractional-factorial structures in strip-plot experiments. These strip-plot experiments are identical in nature to the strip-block experiments, arrangement (c), discussed in Box and Jones (1992). Strip-plot configurations can be applied when the process being investigated is separated into two distinct stages and it is possible to apply the second stage simultaneously to groups of the first-stage product. Miller uses an example involving four washing machines and four dryers in two blocks. Sets of cloth samples are run through the washing machines, and then the samples are divided into groups such that each group contained exactly one sample from each washer. Each group of samples would then be assigned to one of the dryers. The response of interest was the extent of wrinkling.

It is convenient to represent strip-plot structures as rectangular arrays of experimental units in which the levels of one treatment factor (or set of factors) are assigned

to the rows and the levels of a second treatment factor (or set of factors) are assigned to the columns. Table 10 represents the laundry experiment in which each square represents a cloth sample, rows represent sets of samples that were washed together, and columns represent sets of samples that were dried together. The ANOVA table for the laundry example, which is divided into "strata" corresponding to blocks, rows, columns, and units, is given in Table 11. When making inferences about the effects in a particular stratum, the estimate of variation must be based on the residual term for that stratum.

Miller (1997) proposes a method for constructing strip-plot configurations for fractional-factorial designs which consists of three steps:

1. Identify a suitable design for applying row treatments to rows ignoring columns;
2. Identify a suitable design for applying column treatments to columns ignoring rows;
3. Select a suitable fraction of the product of the row and column designs.

The method is applied for two-level designs and then extended to m-level and mixed-level designs. The procedure for two-level designs is presented here; for details on the extended cases, see Miller (1997).

Consider the situation in which a proper fraction of a two-level factorial design is to be run in a strip-plot arrangement using $b = 2^w$ blocks. Each block has $r = 2^M$ rows and $c = 2^m$ columns. Let K and k represent the number of row and column factors, respectively, and define $Q = K - (w + M)$ and $q = k - (w + m)$. Then, the procedure is as follows:

Table 10: Strip-plot Configuration of the Laundry Experiment (from Miller (1997))

	Dryer					Dryer			
Washer	1	2	3	4	Washer	1	2	3	4
1					1				
2					2				
3					3				
4					4				
	Block 1					Block 2			

Table 11: ANOVA Table for the Laundry Example (from Miller (1997))

Strata	Source	df	E(MS)
Block	Blocks	1	$\sigma^2 + 4\sigma_R^2 + 4\sigma_C^2 + 16\sigma_B^2$
Row	W–Washer	3	$\sigma^2 + 4\sigma_R^2 + (8/3)\sum_{j=1}^{4} W_j^2$
	Row Residual	3	$\sigma^2 + 4\sigma_R^2$
Column	D–Dryer	3	$\sigma^2 + 4\sigma_C^2 + (8/3)\sum_{j=1}^{4} D_k^2$
	Column Residual	3	$\sigma^2 + 4\sigma_C^2$
Unit	$W \times D$	9	$\sigma^2 + (1/9)\sum_{j=1}^{4}\sum_{k=1}^{4}[WD]_{jk}^2$
	Unit Residual	9	σ^2

1. Select a row design that consists of a 2^{K-Q} design in b blocks;

2. Select a column design that consists of a 2^{k-q} design in b blocks;

3. Consider the product of the designs in steps 1 and 2 and select a Latin-Square fraction of this product.

The selection of the design in steps 1 and 2 can be made on the basis that the analyses for the row stratum and the column stratum will essentially be the analyses of these designs. The Latin-Square fraction is selected so that the confounding array effects in the unit stratum have desirable properties.

Mee and Bates (1998) consider split-lot experiments involving the etching of silicon wafers. These experiments are performed in steps where a different factor is applied at each step. Thus, there are an equal number of steps and factors. Specifying a split-lot design involves determining the following:

1. the number of process steps with experimentation;

2. the number of factors and their levels at each processing step with experimentation;

3. the subplot size at each processing step;

4. the number of wafers (experimental units) in the entire experiment;

5. a plan that details for each experimental wafer the process subplot at each step.

Mee and Bates emphasize *symmetric* designs, which are designs having the same subplot size at each experimentation step.

The experimental plan in item 5 above will be determined as follows. First, to define b subplots at each step, obtain $b-1$ contrasts for each experimental step. Then assign factors to contrasts within the group intended for their respective processing step. This is done in a way that gives the most information on the interaction effects of interest. The approach is to determine a set of independent contrasts that can be cycled to produce additional sets. The initial set of independent contrasts must be chosen so that the groups of effects remain disjoint. This process and the resulting designs are illustrated for a variety of 64-wafer experiments (see Mee and Bates (1998)). Split-lot designs for three-level factors are also discussed. It should be noted that if there are only two steps, the procedures by Miller (1997) can be applied with one or with many factors at each step.

CHAPTER 3
INCOMPLETE SPLIT-PLOT EXPERIMENTS

The focus of attention in factorial experiments centers on the effects of numerous factors and their interactions. An important class of factorial experiments is the 2^k factorials where each of the k factors is assigned two levels. These experiments are very useful in exploratory investigations as well as optimization problems because they allow a large number of factors and their interactions to be examined.

Since there are only two levels of each factor, they will be denoted as low and high for ease of reference. A treatment combination pertains to a level of each and every factor and will be designated by lower case letters using the following conventions:

- If a factor is at its low level, the corresponding letter is omitted from the treatment designation. Conversely, if a factor is at its high level, the corresponding letter is included.
- When all factors are at their low levels, the treatment will be designated by the symbol (1).

Under this notation, the treatments for a 2^2 factorial experiment in factors P and Q are designated as (1), p, q, and pq. Factors and their effects will be designated by capital letters.

Factorial experiments become large very rapidly so that often a single replicate of the $N = 2^k$ runs requires more resources than are available, even with a moderate

number of factors, k. Even when resources are available, we may not want to estimate all of the $2^k - 1$ factorial effects. As an example, with $k \geq 3$, interactions involving 3 or more factors are generally considered to be negligible or of little importance. Thus, a single replicate of a 2^7 requires 128 experimental units and provides a 64-fold replication of each main effect. Of the 127 effects that can be estimated, only 28 may be of major interest (seven main effects and 21 two-factor interactions).

3.1 Fractional Factorials

Finney (1945) proposed reducing the size of the experiment by using only a fraction of the total number of possible treatment combinations. Such experiments are called fractional factorials. He outlined methods of constructing fractions for 2^k and 3^k experiments. For screening purposes, Plackett and Burman (1946) gave designs for the minimum possible number of experimental units, $N = k + 1$ where N is a multiple of 4, and pointed out their utility in physical and industrial research. Since then, these designs have found many applications, particularly in industrial research and development. Their chief appeal is that they enable a large number of factors, generally 5 or more, to be included in an experiment of practical size so that the investigator can discover quickly which factors have an effect on the response. In this chapter, the discussion will be limited to the case where every factor has only two levels.

A 2^k experiment that is reduced by a factor of 2^{-p} will be called a 2^{k-p} fractional factorial experiment. These experiments have two major problems which can limit their usefulness:

1. Every linear contrast of the treatments estimates more than one effect; hence, each effect is aliased with one or more other effects. This can lead to the misinterpretation of an effect which is not likely to happen with a complete factorial experiment.

2. There is no independent estimate of experimental error.

Despite these limitations, fractional factorial experiments are used in exploratory research and in situations that permit follow-up experiments to be performed. They have been especially useful in industrial research and development where experimental errors tend to be small, the number of factors being investigated is large, and experimentation is sequential. As a tool for exploratory research, fractional factorials provide a means to efficiently evaluate a large number of factors using a relatively small number of experimental units. This allows important factors to be detected and unimportant factors to be screened or discarded rather than committing a large amount of experimental resources on all of the factors.

Effects that are estimated by the same linear combination of treatments are called aliases. Which effects are aliased depends on the factorial effects used to select the treatments. The defining contrast is the effect(s) that is confounded with the constant effect, I. It can be represented as an equation by setting the confounded effect equal to I. The alias chain for an effect is found by forming the generalized interaction of the effect with all terms in the defining contrast. For example, if a 2^{3-1} fraction in factors A, B, and C is run with defining contrast $I = ABC$, then the alias of the main effect A is $A(I) = A^2BC$ which gives $A = BC$ since $A^2 = I$. Therefore, the

alias chains for the main effects, A, B and C are as follows:

$$\begin{aligned} A &= BC \\ B &= AC \\ C &= AB\,. \end{aligned}$$

For a 2^{k-p}, there are $2^p - 1$ effects in the defining contrast. The experimenter can select any p factorial effects to be the defining contrast. The remaining $2^p - p - 1$ factorial effects are automatically determined as being the generalized interactions among the p effects.

Box and Hunter (1961a, 1961b) classified fractional factorial plans by their degree of aliasing of effects. This measure is called the resolution of the plan. The number of letters in the shortest member of a set of defining contrasts determines the design's resolution. Three important resolutions are

1. Resolution III – in which main effects are aliased with two-factor interactions;
2. Resolution IV – in which main effects are aliased with three-factor interactions and two-factor interactions are aliased with other two-factor interactions;
3. Resolution V – where two-factor interactions are aliased with three-factor interactions.

Of course, if all three-factor and higher interactions are negligible, a design with Resolution V is desired because it will allow the estimation of all main effects and two-factor interactions since they are aliased with negligible effects.

3.2 Confounding

Suppose that a 2^k factorial experiment is to be run in blocks. As noted earlier, the main disadvantage of 2^k factorial experiments is their size. Consequently, even for a moderate number of factors, it may not be possible to find blocks with the required number of homogeneous experimental units. When this occurs, it is necessary to use smaller-sized blocks or incomplete block designs.

With an incomplete block design, there must be some loss of information. A balanced incomplete block design, if it exists, distributes this loss equally to all treatments. However, in factorial experiments, it is the main effects and interactions that are important. For most factorial experiments with more than three factors, it is highly unlikely that all effects, especially the higher-order interactions, are important. If some effects can be assumed negligible prior to performing the experiment, then a better procedure for constructing incomplete blocks, originally suggested by Fisher (1926), would be finding arrangements which completely or partially sacrifice the information on these effects so that full information can be obtained on the rest. This is done by forcing the comparisons among the blocks to be identical to the contrasts for the negligible effects. Effects that are estimated by the same linear combination of the treatments are said to be confounded. As a result, it is impossible to determine if the observed difference is due to differences in blocks or due to the factorial effects that are aliased woth the blocks.

Effects selected to be confounded with blocks are called the defining contrasts since they determine which treatments will occur together in a block. These effects are selected by the experimenter and should be effects thought to be negligible since

they are no longer separately estimable. Generally, these effects will be three-factor interactions or higher so that all main effects and two-factor interactions can be estimated.

When the block size of 2^k is reduced by 2^{-p}, each block will contain 2^{k-p} experimental units and each complete replicate will contain 2^p blocks. In this case, it will be necessary to confound $2^p - 1$ effects in each replicate. The experimenter chooses p of these effects with the remaining $2^p - p - 1$ effects being the generalized interactions of the original p effects. When more than one replicate of the 2^{k-p} fractional factorial is performed, two types of confounding are possible:

1. Complete — the same set of effects is confounded in each replicate;

2. Partial — different sets of effects are confounded in different replicates.

Complete confounding is used whenever all information on the confounded effects can be sacrificed. This should only be used when all confounded effects are believed to be negligible. Complete confounding creates no problems with the analysis. It is only necessary to find the effect totals for all unconfounded effects.

There are situations where effects believed to be important must be confounded, for example, when available resources force the use of small block sizes. In these cases, partial confounding is used. Partial confounding means confounding different effects in different replicates so as to allow estimation of all effects. These estimates use only the data from the replicates in which the effect is unconfounded. Thus, there will be greater precision on effects that are unconfounded than on effects that are partially

confounded. While the amount of information is reduced, statistical significance of each effect can be ascertained.

3.3 Confounding in Fractional Factorials

Although only a fraction of the treatments are included in a 2^{k-p} experiment, this number may still be too large for available blocks. As in any factorial experiment, confounding is used to reduce the block size. Confounding an effect in a fractional factorial experiment also confounds all of its aliases.

Consider a 2^{6-1} fractional factorial experiment using the "best" defining contrast for a half-replicate, $I = ABCDEF$. This requires 32 homogeneous experimental units. If these are not available, then blocks of smaller size can be created by confounding additional effects. Suppose blocks of size 16 experimental units are available. To create two blocks of size 16 for the 32 treatments it is necessary to confound one effect. Since $ABCDEF$ was used to define the half-fraction, it would appear logical to select a five-factor interaction, say, $ABCDE$. However, the alias of this interaction or generalized interaction of the effect with $ABCDEF$ is F and will also be confounded with blocks. A better choice is to confound any three-factor interaction since its alias will also be a three-factor interaction. As a result, no information is lost on potentially important effects.

The word "best" should be clarified. It is referring to the design which has the least amount of aliasing among important effects which are usually thought to be main effects and two-factor interactions. If important effects are not aliased with each other, then "best" refers to the design with highest Resolution. Therefore, "best" is

a criterion based on estimability. Throughout this chapter, wherever the phase "best design" is used it will be under the above setting.

3.4 Combining Fractional Factorials and Confounding in Split-Plot Experiments

Splitting the plots or experimental units is possible with any experimental design. The design refers to the assignment of the whole plot and subplot treatments and is selected in order to control the known sources of extraneous variation. Regardless of the choice of design, the subplot treatments can be thought of as being arranged in blocks where the whole plots are the blocks. In each whole plot, if all the subplot treatments can be run, then the situation resembles that of a complete block design as far as the subplot treatments are concerned. However, there are situations where in each whole plot not all of the subplot treatments can be performed so that some form of an incomplete block design must be used. If the subplot treatments result from a 2^k factorial structure, then the methods discussed in the previous sections of this chapter can be applied to reduce the number of subplot treatments in a whole plot.

Consider the situation where both the whole plot treatments and the subplot treatments have a 2^k factorial structure. Assume that the design for the whole plot treatments is a CRD. Suppose, only a fraction of the whole plot treatments are of interest and only a fraction of the subplot treatments can be run for each whole plot. We will consider the situation involving noise factors and design factors. The noise factors will be the whole plot factors. Therefore, the goal of the experiment is to estimate the following:

- main effects for the whole plot factors;
- main effects for the subplot factors;
- two-factor interactions between the whole plot and subplot factors;
- and if possible, two-factor interactions among the subplot factors.

Note that if there were sufficient resources to run all whole plot treatments and subplot treatments, then all four goals would be automatically satisfied. However, in most situations, this is not economically possible. Therefore, we shall try to estimate as many effects as is possible within the restrictions on the resources available.

The idea of confounding effects in order to reduce the number of subplot treatments per whole plot treatment and achieve the second goal has been around for some time. Kempthorne (1952) has a section devoted to confounding in split-plot experiments. Addelman (1964) also discusses ways of accomplishing this. Recently, the use of split-plot experiments in industry has generated renewed interest in confounding. Huang, Chen, and Voelkel (1998) and Bingham and Sitter (1999) discuss minimum-aberration designs for factors with two-levels. This technique helps to improve the estimation problem by raising the resolution concerning the subplot factors, but one must be careful with the whole plot $\times$ subplot interactions. Bisgaard (1999) uses inner and outer arrays, with factors at two-levels, as in robust parameter design and provides the standard errors for various contrasts among the whole plot and subplot factors.

We will use an example to compare the use of confounding in a split-plot experiment. Consider a split-plot experiment with three whole plot factors, A, B, and C,

and three subplot factors, P, Q, and R where all factors have two levels. Suppose only 16 runs are possible among the 64 total number of combinations. There are two ways to allocate the whole plots and subplots for this experiment. We can use four whole plots with each whole plot containing four subplots or we can use eight whole plots with each whole plot containing two subplots. The goal of the experiment is to estimate all six main effects and as many of the nine two-factor interactions between the whole plot and subplot factors as is possible, although it is believed that some two-factor interactions among the subplot factors might be significant. To conserve space in the tables, the confounding structure or alias chains will be given only up to order two. Therefore, if there is a blank space in the alias table, it means that the effect is aliased with interactions of order higher than two.

First, suppose that the experimenter ignores the split-plot structure by considering the factors as a 2^6 factorial in a completely randomized design. Actually, the experimenter would use a 2^{6-2} fractional factorial design to obtain the 16 runs. The best defining contrast is

$$I = ABCP = CPQR = ABQR,$$

which has Resolution IV. The layout is given in Table 12 and the alias chains are given in Table 13. All main effects can be estimated, but two-factor interactions are aliased with each other. Even if we assume that all two-factor interactions among A, B, and C and all two-factor interactions among P, Q, and R are negligible, there is still a problem since AQ is aliased with BR and AR is aliased with BQ. In other words, some of the interactions we are interested in are aliased with each other.

Table 12: Design Layout for 2^{6-2} With Defining Contrast $I = ABCP = CPQR = ABQR$

abcp	*ab*	*cp*	*abcpqr*
acr	*acq*	*cpqr*	*abqr*
bcq	*bcr*	*bpr*	*bpq*
apq	*qr*	*apr*	(1)

Table 13: Alias Structure for 2^{6-2}

A	$=$	
B	$=$	
C	$=$	
P	$=$	
Q	$=$	
R	$=$	
AB	$=$	$CP + QR$
AC	$=$	BP
AP	$=$	BC
AQ	$=$	BR
AR	$=$	BQ
CQ	$=$	PR
CR	$=$	PQ

Table 14: Design Layout for the Combined $2^{3-1} \times 2^{3-1}$
With Defining Contrast $I = ABC = PQR = ABCPQR$

a	b	c	abc
p	p	p	p
q	q	q	q
r	r	r	r
pqr	pqr	pqr	pqr

A second method, incorporating the split-plot nature and using four whole plots, is to consider reducing the whole plot factors and subplot factors separately using fractional factorials. A 2^{3-1} fractional factorial with defining contrast $I = ABC$ will be used for selecting the whole plot treatments and combined with a 2^{3-1} with defining contrast $I = PQR$ in selecting the subplot treatments (see Table 14). The overall defining contrast for the experiment is

$$I = ABC = PQR = ABCPQR\,,$$

and the alias structure is shown in Table 15. Once we consider the split-plot structure, the best we can do at the whole plot level is a Resolution III design. This method provides a good design for estimating the two-factor interactions between the whole plot and subplot factors. However, we must assume that the two-factor interactions among the subplot factors are negligible in order to estimate the main effects for the subplot factors.

Method three uses split-plot confounding and four whole plots. At the whole plot level, a 2^{3-1} fractional factorial with defining contrast $I = ABC$ is used. Then, the three-factor interaction, PQR, is confounded with factor C to reduce the eight

Table 15: Alias Structure for $2^{3-1} \times 2^{3-1}$

A	=	BC
B	=	AC
C	=	AB
P	=	QR
Q	=	PR
R	=	PQ
AP	=	
AQ	=	
AR	=	
BP	=	
BQ	=	
BR	=	
CP	=	
CQ	=	
CR	=	

subplot treatments to four per whole plot (see Table 16). The idea is to put the positive fraction of PQR wherever C is positive and the negative fraction wherever C is negative. The overall defining contrast is given by

$$I = ABC = CPQR = ABPQR$$

with the alias structure provided in Table 17. This design is better than the second design in terms of aliasing of the main effects for the subplot factors, but cannot estimate all nine whole plot by subplot factor interactions without assuming that PQ, PR, and QR are negligible. If, on the other hand, it is reasonable to assume that the whole plot factor C will not interact with any of the subplot factors, then PQ, PR, QR, the main effects for subplot factors and the remaining six whole plot by subplot

Table 16: Design Layout for Split-Plot Confounding
With Defining Contrast $I = ABC = CPQR = ABPQR$

a	b	c	abc
(1)	(1)	p	p
pq	pq	q	q
pr	pr	r	r
qr	qr	pqr	pqr

factor interactions can be estimated using this design. Also, on a consulting level, some experimenters would feel more comfortable with this design since it uses all 8 subplot treatments.

The fourth method uses eight whole plots and split-plot confounding. Since there are eight whole plots, the complete 2^3 factorial can be used for the whole plot factors. However, we must now reduce the number of subplots to two per whole plot. This implies that we must confound two members in the defining contrast and their generalized interaction completes the defining contrast. Using split-plot confounding, the defining contrast is

$$I = ABPQ = ACQR = BCPR,$$

with the layout given in Table 18 and the alias structure given in Table 19. This design is good for estimating main effects but has some serious deficiencies with interactions.

One possible problem with designs that use eight whole plots is cost. If the whole plot factors are costly to change, then using eight whole plots as opposed to four might be impractical. Another problem with designs using eight whole plots is the breakdown of the degrees of freedom. There are 7 df for the whole plot design and

Table 17: Alias Structure for Split-Plot Confounding

A	$=$	BC
B	$=$	AC
C	$=$	AB
P	$=$	
Q	$=$	
R	$=$	
AP	$=$	
AQ	$=$	
AR	$=$	
BP	$=$	
BQ	$=$	
BR	$=$	
CP	$=$	QR
CQ	$=$	PR
CR	$=$	PQ

Table 18: Design Layout for Split-Plot Confounding in 8 Whole plots With $I = ABPQ = ACQR = BCPR$

(1)	a	b	ab	c	ac	bc	abc
AB^+ AC^+	AB^- AC^-	AB^- AC^+	AB^+ AC^-	AB^+ AC^-	AB^- AC^+	AB^- AC^-	AB^+ AC^+
PQ^+ QR^+	PQ^- QR^-	PQ^- QR^+	PQ^+ QR^-	PQ^+ QR^-	PQ^- QR^+	PQ^- QR^-	PQ^+ QR^+
pqr (1)	pr q	qr p	pq r	pq r	qr p	pr q	pqr (1)

Table 19: Alias Structure for Split-Plot Confounding
For 8 Whole Plots With $I = ABPQ = ACQR = BCPR$

A	$=$	
B	$=$	
C	$=$	
AB	$=$	PQ
AC	$=$	QR
BC	$=$	PR
P	$=$	
Q	$=$	
R	$=$	
AP	$=$	BQ
AQ	$=$	$BP + CR$
AR	$=$	CQ
BR	$=$	CP

only 8 df left for the subplot factors and whole plot $\times$ subplot factor interactions. Therefore, at the subplot level there are only enough df to estimate either three main effects and five interactions or eight interactions. This may not be sufficient to estimate all the effects of interest.

3.5 Discussion of Minimum-Aberration Split-Plot Designs

In split-plot designs using some sort of confounding, there is a concept of partial resolution. The partial resolution of the whole plots refers to the resolution of terms in the defining contrast involving only whole plot factors. The partial resolution of the subplot factors refers to the resolution of terms in the defining contrast involving either only subplot factors or both whole plot and subplot factors. Recall that the definition of minimum-aberration is the design that has smallest number of words in

the defining contrast with the fewest letters. Therefore, it is looking at the overall resolution of the design and not the partial resolution.

Huang, Chen and Voelkel (1998) and Bingham and Sitter (1999) have tabled minimum-aberration (MA) designs for 16 and 32 runs for up to 10 factors. When a design is needed that fits in these restrictions, one can simply look up the appropriate design in these tables. However, the MA designs in these tables do not take into account other design issues such as which effects are the most important to estimate. This concept seems to be overlooked in the literature. For example, suppose the whole plot factors are noise factors and only their main effects are of interest. Now, further suppose that the two-factor whole plot by subplot interactions are the most important effects to estimate (which is the case in many experiments). Then, it is better to fractionate the whole plot treatments and subplot treatments separately since this would alias the two-factor interactions of interest with higher order interactions. Note, this design would not be the MA design since the partial resolution of the whole plot factors would be too low.

Another concern with MA designs is the allocation of the runs. Consider the MA designs for 16 runs involving combinations of 2, 3, and 4 whole plot and subplot factors. With the exception of the cases involving 2 whole plot factors with 3 or 4 subplot factors, all the other MA designs use eight whole plots with two subplots per whole plot. This raises several concerns.

1. Typically in industrial experiments, the whole plot factors are hard-to-change or costly-to-change factors. If they are hard to change, then it would make more sense to only change them four times as opposed to eight. Also, if changing

these factors is expensive, then again changing them four times seems more reasonable.

2. It is not an efficient allocation of the degrees of freedom. Using eight whole plots with two subplots per whole plot gives 7 df for whole plot factors and 8 df for subplot factors and whole plot × subplot factor interactions. This allocates a disproportionate number of degrees of freedom to the whole plot factors. In contrast, using four whole plots with four subplots per whole plot gives 3 df for whole plot factors and 12 df for terms involving subplot factors.

3. Using two subplots per whole plot is similar to using blocks of size two in a block design which is not generally recommended.

MA designs are in general "good" designs, however, for split-plot experiments they are based purely on the overall resolution of the design instead of partial resolution. Also, they only use split-plot confounding to reduce the size of the experiment and are not motivated by any other concerns such as those mentioned above.

3.6 Adding Runs to Improve Estimation

With the concerns of the previous section in mind, mainly the allocation of degrees of freedom, we will focus our attention on 16 run designs that use four whole plots with four subplots per whole plot. Within this allocation of the resources, the best design is found for the two types of confounding discussed in the example in section 3.4. These are split-plot confounding and fractionating of the whole plot and subplot factors separately (called the Cartesian product design in Bisgaard (1999)). The best

design is found using the "minimum-aberration" (MA) criterion, but this differs from just using MA because we are restricted to using four whole plot treatments with four subplot treatments per whole plot. Therefore, the best resolution is desired within this restricted setting. With the exception of the cases involving 2 whole plot factors with 3 or 4 subplot factors, these designs will not be the overall MA design.

Once the sixteen run design is found, eight additional runs are considered in order to break some of the alias chains. Along with breaking some of the alias chains, extra degrees of freedom are now available in order to estimate additional effects. The result is a 24 run design which we feel is a nice compromise between the 16 and 32 run designs presented in Huang, Chen and Voelkel (1998) and Bingham and Sitter (1999). Which eight treatments should be added is the question to be answered next, but first we briefly discuss foldover designs.

The concept of a foldover design was introduced in Box and Hunter (1961b). Suppose an experiment involving k factors each at two levels is to be performed and an initial Resolution III fractional factorial design is used. One way to do a the foldover is to repeat the initial design and change the levels of one of the factors while leaving the levels of the other factors unchanged. This allows the estimation of all the interactions that contain the folded factor but doubles the size of the experiment. A related idea is that of semifolding which folds only the points that are at the high level of a factor (or the low level). The addition of the new points breaks certain alias chains and allows estimates of interactions involving the factor that is semifolded to be calculated while adding only half as many points as a complete foldover design. In the rest of this chapter, we apply semifolding to split-plot experiments.

In most of the cases studied here, the eight additional points are added to the initial 16 run design by semifolding on either one or two subplot factors which results in a 24 point design consisting of four whole plots with six subplots per whole plot. These designs will have 3 df for the whole plot treatments and 20 df for the subplot treatments. The initial 16 point design is balanced over the subplot factors—each factor has the same number of high and low levels present—which allows for the effects to estimated with equal precision. It is desired to preserve this balance of the subplot factors in the 24 point design as well as maintain the same number of subplot treatments per whole plot. Therefore, in half of the whole plots the semifolding is on the high level of a subplot factor while in the other half the semifolding is on the low level of that factor.

In some cases, it is necessary to fold on a whole plot factor in order to estimate the main effects of the whole plot factors. In these cases, two whole plots are added so that the 24 point design consists of six whole plots with four subplots per whole plot. These designs will have 5 df for the whole plot treatments and 18 df for the subplot treatments. All nine cases involving 2, 3, and 4 whole plot and subplot factors are considered. However, two cases do not need to be improved upon.

1. Two whole plot factors and two subplot factors: the 16 points represent the full factorial. Since no fractionating or confounding is needed, there is nothing to improve upon.

2. Two whole plot factors and three subplot factors: in this case, the MA design presented in Huang, Chen and Voelkel (1998) is the best design possible and

allows estimates all of the main effects and all of the two-factor interactions.

For all situations involving less than four whole plot factors, a general method can be used to construct 24 run designs. First, construct a 16 run design that uses four whole plots with four subplots in each whole plot.

- If there are two whole plot factors, then use the complete factorial in the whole plot factors.
- If there are three whole plot factors, then use one of the two half fractions found using the defining contrast $I = ABC$.

To complete the 16 run design use either split-plot confounding or a separate fractional factorial in the subplot factors to decide which subplot treatments will appear in each whole plot. After the 16 run design is selected on, use semifolding to obtain two extra subplot treatments for each whole plot treatment. The semifolding is, for the most part, done on two subplot factors. In two of the whole plots, the subplot treatments are folded on one factor (the high level of the factor in one whole plot and the low level of the factor in the other whole plot). In the remaining two whole plots, the subplot treatments are folded on a different subplot factor (again, on the high level in one whole plot and the low level in the other whole plot). In the special case of three subplot factors, the semifolding is done on just one factor since there is only one alias chain in the defining contrast.

When there are four or more whole plot factors, using 4 whole plots results in insufficient degrees of freedom to estimate the main effects of the whole plot factors.

Therefore, the additional eight runs will be added in the form of two extra whole plots. This leads to a 24 run design with 6 whole plots with 4 subplots per whole plot. The whole plot treatments used in the two additional whole plots are found by semifolding on a whole plot factor. However, which subplot treatments should be used in the two additional whole plots is case specific. To illustrate how to apply these methods, the seven remaining cases involving 2, 3 and 4 whole plot and subplot factors will be presented. For all the cases, tables which show the designs in highs and lows for each factor are given in Appendix A.

For these designs, the analysis could use one normal probability plot for the whole plot effects and a separate plot for the effects involving the subplot factors and the interactions between whole plot and subplot factors. One assumption of a normal probability plot is that the effects are independent. This is not the case here. However, it will be shown later in this chapter that the correlations are low (near zero) and that a normal probability plot is therefore valid. In some cases there are degrees of freedom left at the subplot level so that if desired they can be used to estimate the an error variance. It should be reiterated that the goal of the experiment is to estimate as well as test for significance the main effects for the whole plot factors, the two-factor interactions between whole plot and subplot factors, the main effects of the subplot factors, and if possible two-factor interactions among the subplot factors.

2 WP Factors (A, B) and 4 SP Factors (P, Q, R, S)

To obtain a 16 point design under this situation, only the subplot treatments need to be fractionated or confounded. First, consider fractionating the subplot treatments.

The defining contrast is $I = PQR = QRS = PS$ which is resolution II. Alias chains involving both P and S need to be broken. Therefore, the additional eight points are obtained by semifolding on high and low P in two whole plots and on high and low S in the other two. The 24 point design is shown in Table 20. The chains are almost completely broken. Only two of the three interactions, BP, BS, and PS are estimable. If it can be assumed that PS is negligible, then everything else is estimable. It is not unreasonable to believe that with four factors one of the two-factor interactions is negligible and the experimenter should be able to help determine which interaction is most likely to be negligible.

Next, consider split-plot confounding. The defining contrast is $I = APQR = BQRS = ABPS$ which is resolution IV. This design is the MA design given in Huang, Chen and Voelkel (1998). Again, alias chains involving both P and S need to be broken. Therefore, the additional eight points are obtained by semifolding on high and low P in two whole plots and on high and low S in the other two. The 24 point design is shown in Table 21. All of the two-factor interactions between whole plot and subplot factors can be estimated except BS which is aliased with QR. If QR is assumed to be negligible, then BS can be estimated. Most of the two-factor interactions among the subplot factors are aliased with each other. However, with the 24 point design we can estimate three of the two-factor interactions between the subplot factors without making any assumptions about negligibility, which is an improvement over the MA design.

Table 20: 24 Point Design for the Case of 2 WP Factors and 4 SP Factors Using the Same Fraction [HP(HS)-denotes high $P(S)$ and LP(LS)-denotes low $P(S)$]

a	b	ab	(1)
q	q	q	q
r	r	r	r
ps	ps	ps	ps
$pqrs$	$pqrs$	$pqrs$	$pqrs$

Fold on

HP	LP	HS	LS
s	pq	p	qs
qrs	pr	pqr	rs

Table 21: 24 Point Design for the Case of 2 WP Factors and 4 SP Factors Using Split-Plot Confounding [HP(HS)-denotes high $P(S)$ and LP(LS)-denotes low $P(S)$]

a	b	ab	(1)
p	s	q	qr
pqr	pq	r	pqs
qs	pr	ps	prs
rs	qrs	$pqrs$	(1)

Fold on

HP	LP	HS	LS
(1)	ps	p	qrs
qr	$pqrs$	pqr	s

Table 22: 16 Point Design for the Case of 3 WP Factors and 2 SP Factors Using a Fraction Factorial of the Whole Plot Factors

a	b	c	abc
p	p	p	p
q	q	q	q
pq	pq	pq	pq
(1)	(1)	(1)	(1)

3 WP Factors (A, B, C) and 2 SP Factors (P, Q)

In this case, only the whole plot factors need to be fractionated. Since nothing needs to be done to the subplot factors, there is only one 16 point design. The defining contrast is $I = ABC$ which is Resolution III. Since the design estimates everything set out in the goal of the experiment, no points need to be added to this design. However, note that this is not the MA design which is run using eight whole plots with 2 subplots per whole plot. The 16 point design in four whole plots with four subplots per whole plot is shown in Table 22.

3 WP Factors (A, B, C) and 3 SP Factors (P, Q, R)

To obtain a 16 point design in this situation, both the whole plot and subplot treatments need to be fractionated or confounded. First, consider fractionating the whole plot and subplot treatments separately. The defining contrast is $I = ABC = PQR = ABCPQR$ which is resolution III. The two-factor interactions between whole plot and subplot factors are already estimable. Therefore, there is only one alias chain that needs to be broken, and that is associated with PQR. The additional

Table 23: 24 Point Design for the Case of 3 WP Factors and 3 SP Factors Using the Same Fraction [HP-denotes high P and LP-denotes low P]

a	b	c	abc
p	p	p	p
q	q	q	q
r	r	r	r
pqr	pqr	pqr	pqr

Fold on

HP	LP	LP	HP
(1)	pq	pq	(1)
qr	pr	pr	qr

eight points are obtained by semifolding on factor P. The 24 points design is shown in Table 23. The chain has been broken and now P, Q, R, PQ, PR, and QR are all estimable. There are 5 df left over for a subplot error term.

Next, consider split-plot confounding. The defining contrast is $I = ABC = ABPQR = CPQR$ which is also resolution III. The two-factor interactions between whole plot factors A and B and the subplot factors are already estimable. Therefore, the only alias chain that needs to be broken is $CPQR$. The additional eight points are obtained by semifolding on factor P while being careful to fold both high and low P where C is high and where C is low. The 24 point design is shown in Table 24. The chain has been broken and now all of the effects of interest including the two-factor interactions among the subplot factors are estimable. There are 5 df left over for a subplot error term.

Table 24: 24 Point Design for the Case of 3 WP Factors and 3 SP Factors Using Split-Plot Confounding [HP-denotes high P and LP-denotes low P]

a	b	c	abc
pq	pq	p	p
pr	pr	q	q
qr	qr	r	r
(1)	(1)	pqr	pqr

Fold on

HP	LP	LP	HP
q	pqr	pq	(1)
r	p	pr	qr

3 WP Factors (A, B, C) and 4 SP Factors (P, Q, R, S)

To obtain a 16 point design in this situation, both the whole plot and subplot treatments need to be fractionated or confounded. First, consider fractionating the whole plot and subplot treatments separately. The defining contrast is $I = ABC = PQR = QRS = PS = ABCPQR = ABCQRS = ABCPS$ which is Resolution II. In order to estimate the subplot factor main effects and possibly the two-factor interactions among the subplot factors, the two chains PQR and QRS with resulting chain PS need to be broken. The additional eight points are obtained by semifolding on both factors P and S. The 24 points design is shown in Table 25. The chains are almost completely broken. Two resulting chains $AP = PS$ and $AS = PS$ are left. The aliasing here means that the sum of AP and AS equals PS. Therefore, the model can accomodate the fitting of any two of the three factors. So for example, if PS is assumed negligible, then the effects of AS and AP are estimable. There are 4 df left that can be used as an error term or used to estimate PQ, PR, QS and RS.

Table 25: 24 Point Design for the Case of 3 WP Factors and 4 SP Factors Using the Same Fraction [HP(HS)-denotes high $P(S)$ and LP(LS)-denotes low $P(S)$]

a	b	c	abc
q	q	q	q
r	r	r	r
ps	ps	ps	ps
$pqrs$	$pqrs$	$pqrs$	$pqrs$

Fold on

HP	LP	HS	LS
s	pq	p	qs
qrs	pr	pqr	rs

Next, consider split-plot confounding. The defining contrast is $I = ABC = BCPQR = ACQRS = ABPS = APQR = BQRS = CPS$ which is resolution III. Not much of anything is estimable free of two-factor interactions. Again, the additional eight points are obtained by semifolding on factors P and S. The 24 point design is shown in Table 26. Again, the chains are almost completely broken. Three resulting chains $C = PS$, $AP = PS$ and $BS = PS$ are left. The aliasing here means that the sum of C, AP, and BS equals PS. Therefore, the model can accomodate the fitting of any three of the four factors. So for example, assuming PS is negligible allows for C, AP and BS to be estimated. Also, any two of the remaining five two-factor interactions among the subplot factors can be estimated.

4 WP Factors (A, B, C, D) and 2 SP Factors (P, Q)

In this case, only the whole plot treatments need to be fractionated. Note, with four whole plot factors there are only 3 df for whole plot factor effects. Hence, whole

Table 26: 24 Point Design for the Case of 3 WP Factors and 4 SP Factors Using Split-Plot Confounding [HP(HS)-denotes high $P(S)$ and LP(LS)-denotes low $P(S)$]

a	b	c	abc
p	s	qr	q
pqr	pq	pqs	r
qs	pr	prs	ps
rs	qrs	(1)	$pqrs$

Fold on

HP	LP	LS	HS
(1)	ps	qrs	p
qr	$pqrs$	s	pqr

Table 27: 24 Point Design for the Case of 4 WP Factors and 2 SP Factors Using a Fractional Factorial of the Whole Plot Factors

b	c	ad	$abcd$
p	p	p	p
q	q	q	q
pq	pq	pq	pq
(1)	(1)	(1)	(1)

Fold on A

d	bcd
p	p
q	q
pq	pq
(1)	(1)

plots will need to be added for all cases involving four whole plot factors. Since nothing needs to be done to the subplot factors, there is only one 16 point design. The defining contrast is $I = ABC = BCD = AD$ which is resolution II. The additional whole plots are obtained by semifolding on factor A. The 24 point design is shown in Table 27. The chains are broken and everything is estimable.

4 WP Factors (A, B, C, D) and 3 SP Factors (P, Q, R)

To obtain a 16 point design in this situation, both the whole plot and subplot treatments need to be fractionated or confounded. First, consider fractionating the whole plot and subplot treatments separately. The defining contrast is $I = ABC = BCD = AD = PQR = ABCPQR = BCDPQR = ADPQR$ which is resolution II. The additional whole plot treatments are obtained by semifolding on factor A. The positive fraction, $I = PQR$, is run in one whole plot while the negative fraction, $I = -PQR$, is run in the other whole plot. The negative fraction can be thought of as semifolding on any subplot factor and placing all of the points in one whole plot instead of two as was done in all the cases up until now. The 24 point design is shown in Table 28. The chains are broken and everything is estimable.

Next, consider split-plot confounding. The defining contrast is $I = ABC = BCD = AD = CPQR = ABPQR = BDPQR = ACDPQR$ which is resolution II. Besides breaking chains among the whole plot factors, the chain, $CPQR$ needs to be broken. The additional whole plot treatments are obtained by semifolding on factor A. Again, the positive fraction, $I = PQR$, is run in one whole plot with the negative fraction, $I = -PQR$, is run in the other whole plot. Again, this can be thought of as semifolding each fraction on any subplot factor and placing all four points in the same whole plot. The 24 point design is shown in Table 29. The chains are broken and everything is estimable.

Table 28: 24 Point Design for the Case of 4 WP Factors and 3 SP Factors Using the Same Fraction

b	c	ad	$abcd$
p	p	p	p
q	q	q	q
r	r	r	r
pqr	pqr	pqr	pqr

Fold on A

d	bcd
p	pq
q	pr
r	qr
pqr	(1)

Table 29: 24 Point Design for the Case of 4 WP Factors and 3 SP Factors Using Split-Plot Confounding

b	c	ad	$abcd$
pq	p	pq	p
pr	q	pr	q
qr	r	qr	r
(1)	pqr	(1)	pqr

Fold on A

d	bcd
p	pq
q	pr
r	qr
pqr	(1)

4 WP Factors (A, B, C, D) and 4 SP Factors (P, Q, R, S)

As the number of both whole plot factors and subplot factors increases, it becomes impossible to break all of the relationships and estimate all of the important effects. Therefore, some effects will need to be assumed negligible. Also, in the case of four whole plot factors and four subplot factors, there is insufficient degrees of freedom to estimate the four subplot factor main effects and the sixteen two-factor whole plot by subplot interactions. Thus, some effects cannot be estimated anyway. Assuming these effects to be negligible enables the estimation of the remaining effects.

To obtain a 16 point design in this situation, both the whole plot and subplot treatments need to be fractionated. First, consider fractionating the whole plot and subplot treatments separately. The defining contrast is $I = ABC = BCD = AD = PQR = QRS = PS = ABCPQR = ABCQRS = ABCPS = BCDPQR = BCDQRS = BCDPS = ADPQR = ADQRS = ADPS$ which is resolution II. Besides breaking chains among the whole plot factors, the chains, $ADPS$ and PS need to be broken. The additional whole plot treatments are obtained by semifolding on factor A. The subplots are semifolded on factor P in one whole plot and factor S in another whole plot. The 24 point design is shown in Table 30.

Next, consider split-plot confounding. The defining contrast is $I = ABC = BCD = AD = ACPQR = BDQRS = ABCDPS = BPQR = ABDPQR = CDPQR = ACDQRS = CQRS = ABQRS = DPS = APS = BCPS$ which is resolution II. The additional whole plot treatments are obtained by semifolding on factor A. Again, the subplot factors are semifolded on P and S. Care must be taken

when choosing which whole plots the subplot factors are semifolded. Otherwise, the same treatment combinations will occur in both additional whole plots. This occurs when the semifolding uses the whole plots containing the subplot treatments defined by PQR^+, QRS^+ and PQR^-, QRS^- or PQR^+, QRS^- and PQR^-, QRS^+. Any other combination is fine. In this section, P is semifolded in the whole plot containing whole treatment c (PQR^-, QRS^+) and S is semifolded in the whole plot containing whole plot treatment *abcd* (PQR^+, QRS^+). The 24 points design is shown in Table 31.

Most of the chains are broken but some of the two-factor interactions among the subplot factors are aliased with each other. Also, four of the sixteen two-factor interactions between whole plot and subplot factors must be assumed negligible. These terms are AS, CS, DS, and DP. This is fairly nice since three of these terms involve subplot factor S. Therefore, if it is believed that one of the subplot factors is unlikely to interact with the whole plot factors, these terms or effects could be assumed negligible. This does not seem unreasonable. Now the 18 subplot df are partitioned into 4 df for the subplot factor main effects, 12 df for the whole plot by subplot interactions, and 2 for two-factor interactions among subplot factors (these two effects can be any pair except PQ and QS or PR and RS).

3.7 An Example

To illustrate how an experiment could be carried out and analyzed, an example is presented. The example, from Taguchi (1987), involves the study of a wool washing and carding process. The original experiment used a $2^{13-9} \times 2^{3-1}$ inner and outer

Table 30: 24 Point Design for the Case of 4 WP Factors and 4 SP Factors Using the Same Fraction

b	c	ad	$abcd$
q	q	q	q
r	r	r	r
ps	ps	ps	ps
$pqrs$	$pqrs$	$pqrs$	$pqrs$

Fold on A

d	bcd
pq	qs
pr	rs
s	p
qrs	pqr

Table 31: 24 Point Design for the Case of 4 WP Factors and 4 SP Factors Using Split-Plot Confounding

b	c	ad	$abcd$
p	qrs	(1)	q
pqr	s	qr	r
qs	pq	pqs	ps
rs	pr	prs	$pqrs$

Fold on A

d	bcd
$pqrs$	qs
ps	rs
q	p
r	pqr

array. For our purposes, we shall only consider the first three factors in the inner array along with the three factors in the outer array. The outer array will make up the whole plot factors while the inner array will have the subplot factors. The original experiment was not run using restricted randomization, but it will be assumed that it was in order to present the analysis.

The 24 run designs discussed earlier can be utilized. For this example, the design given in Table 23 with separate fractions at the whole plot and subplot levels is used. The only difference is that the Taguchi example uses the negative fraction of the whole plot factors instead of the positive fraction. To correspond with the factor names in the Taguchi example, let X, Y and Z be the whole plot factors and A, B and C be the subplot factors. The design along with the responses is shown in Table 32.

The analysis involves fitting the 18-term model involving the main effects of the whole plot factors, the main effects of the subplot factors, the two-factor interactions among subplot factors, and two−factor interactions between whole plot and subplot factors. This leaves five degrees of freedom for a subplot error term. Table 33 gives the estimated effects and t-tests. The tests for the three whole plot factor main effects are not correct and should be ignored. From Table 33, there appears to be an effect due to the interaction of factors B and C. Since there are only 5 df for error, one might choose to use a normal probability plot to investigate at the subplot level. The design is not completely balanced or orthogonal which leads to some effects having one standard error and others having a different standard error. Therefore, instead of just plotting the effects, the effects are divided by their standard error and then plotted. The plot is shown in Figure 1 and gives BC and ZC as significant effects.

Table 32: 24 Point Design for the Example

X	Y	Z	A	B	C	Response
−	−	−	+	−	−	19.0
			−	+	−	22.5
			−	−	+	26.0
			+	+	+	21.5
			−	−	−	20.0
			−	+	+	18.5
+	−	+	+	−	−	16.0
			−	+	−	21.0
			−	−	+	20.0
			+	+	+	16.0
			+	+	−	24.0
			+	−	+	23.0
−	+	+	+	−	−	17.5
			−	+	−	21.0
			−	−	+	22.0
			+	+	+	22.0
			+	+	−	21.0
			+	−	+	19.0
+	+	−	+	−	−	19.0
			−	+	−	22.5
			−	−	+	26.5
			+	+	+	23.0
			−	−	−	23.0
			−	+	+	26.5

Table 33: Effects Table for the Example

Term	Effect	Coeff	Std Error	t-value	P-value
Constant		21.266	0.4288	49.60	0.000
X	−0.937	−0.469	0.4951	−0.95	0.387†
Y	1.406	0.703	0.4288	1.64	0.162†
Z	−2.156	−1.078	0.4288	−2.51	0.054†
A	1.000	0.500	0.4951	1.01	0.359
B	−0.812	−0.406	0.4951	−0.82	0.449
C	0.937	0.469	0.4951	0.95	0.387
$X*A$	0.031	0.016	0.4288	0.04	0.972
$X*B$	−0.187	−0.094	0.4951	−0.19	0.857
$X*C$	−0.563	−0.281	0.4951	−0.57	0.595
$Y*A$	−0.094	−0.047	0.4288	−0.11	0.917
$Y*B$	0.312	0.156	0.4288	0.36	0.730
$Y*C$	0.875	0.438	0.4288	1.02	0.354
$Z*A$	0.344	0.172	0.4288	0.40	0.705
$Z*B$	0.500	0.250	0.4288	0.58	0.585
$Z*C$	−1.312	−0.656	0.4288	−1.53	0.186
$A*B$	1.375	0.687	0.4951	1.39	0.224
$A*C$	1.250	0.625	0.4951	1.26	0.262
$B*C$	−4.438	−2.219	0.4951	−4.48	0.007

† Not Valid Tests

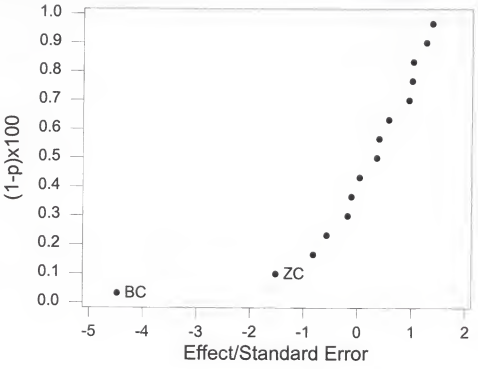


Figure 1: Normal Probability Plot for the Example

3.8 Summary

The main goal of this chapter is to understand some of the complications involved with using the two types of confounding in split-plot experiments. If the design is chosen using the MA criterion, then designs constructed by combining the fractional factorial of the whole plot treatments with a fractional factorial of the subplot treatments are not considered. This is because their partial resolution is too low. However, it has been shown that depending on what is to be estimated, it may not be wise to eliminate such designs. Also, for 16 run designs the MA criterion tends to select designs with eight whole plots and two subplots per whole plot. These designs do not take into consideration the possible increase in the cost of experimentation or of factors whose levels are hard to change.

We have presented a 24 run design which is a compromise between the 16 run and 32 run designs. We begin with a 16 run design using four whole plots with four subplots per whole plot. Such a strategy accommodates hard-to-change factors as well as cost considerations involving the experiment. Then, eight additional runs are added using semifolding of one or two factors. These runs preserve the balance of high and low levels of each subplot factor as well as maintain the same number of subplots for each whole plot. Except for the cases involving four whole plot factors, the additional runs are at the subplot level. Thus, as long as adding the runs is feasible, it should not be too costly. Also, the extra runs allow for additional effects to be estimated and/or add degrees of freedom for estimating the subplot error variance. Though designs using split-plot confounding and separate fractions differ in what they can estimate in 16 runs, there is not much difference in estimability when using

Table 34: Condition Numbers for the Various Cases

# of WP Factors	# of SP Factors	Type of Confounding	Condition Number, κ
2	4	Same Fraction	4.05
2	4	Split-Plot Confounding	3.36
3	3	Same Fraction	2.00
3	3	Split-Plot Confounding	2.00
3	4	Same Fraction	4.05
3	4	Split-Plot Confounding	3.36
4		Neither	2.00
4	3	Same Fraction	5.83
4	3	Split-Plot Confounding	2.00
4	4	Same Fraction	5.83
4	4	Split-Plot Confounding	3.36

the 24 run designs.

The phrase "the chains are broken" used throughout this chapter does not mean that all of the effects are no longer aliased. Sometimes the effects are aliased at a lower degree than unity (completely aliased). Therefore, there is some degree of collinearity between the effects. To measure the magnitude of this collinearity, the condition number

$$\kappa = \sqrt{\frac{\text{Largest Eigenvalue of } (X'X)}{\text{Smallest Eigenvalue of } (X'X)}}$$

is calculated for each case (see Table 34). Many textbooks declare collinearity to be a problem if $\kappa \geq 30$. It is seen from Table 34 that collinearity does not seem to be a problem for the cases considered in this chapter. The variance inflation factors (VIF)

are also calculated. They are not reported here, but none of the VIF's exceeds 3. This confirms that the collinearity is mild.

The cases used in this chapter are chosen because they cover a wide range of industrial applications. The methods described in this chapter can be applied to experiments involving even more than four factors at either or both the whole plot and subplot level. However, as is seen with the case involving four whole plot factors and four subplot factors, not all chains can be broken. However, the additional points should result in a design which estimates more effects with less assumptions on negligibility than a 16 point design. If the cost and time of adding the eight points is acceptable and the goals of the experiment are those discussed in this chapter, then the 24 run designs can improve the estimability of important effects.

CHAPTER 4
A NEW MODEL AND CLASS OF DESIGNS FOR MIXTURE EXPERIMENTS WITH PROCESS VARIABLES

Experiments that involve the blending of two or more components to produce high quality products are known as mixture experiments. The quality of the end product depends on the relative proportions of the components in the mixture. For example, suppose we wish to study the flavor of a fruit punch consisting of juices from apples, pineapples, and oranges. The flavor of the punch depends on the relative proportion of the juices in the blend.

Consider a mixture experiment consisting of q components. Let x_i, for $i = 1, 2, \ldots, q$, represent the fractional proportion contributed by component i. Then the proportions must satisfy the following constraints

$$0 \leq x_i \leq 1\,, \quad \sum_{i=1}^{q} x_i = 1\,,$$

and the experimental region is a $(q-1)$-dimensional simplex, S_q. For $q = 3$, S_3 is an equilateral triangle and for $q = 4$, S_4 is a tetrahedron. Typically, the blends used in a mixture experiment are the vertices or single-component blends, the midpoints of the edges, centroids of faces, etc., and the centroid of the simplex.

In some mixture experiments, the quality of the product depends not only on the proportions of the components in the blend, but also on the processing conditions. Process variables are factors that do not form any portion of the mixture but whose levels, when changed, could affect the blending properties of the components. Cornell

(1990) discusses an experiment involving fish patties. The texture of the fish patties depends not only on the proportions of three fish species that are blended but also on three process variables which are cooking temperature, cooking time and deep fat frying time.

A concern with mixture experiments involving process variables is that the size of the experiment increases rapidly as the number of process variables, n increases. In the fruit punch or fish patty examples above, it may not be necessary to limit the size of the experiment. However, in most industrial experiments, cost and time do impose restrictions on the number of runs permitted. Therefore, a design strategy that uses fewer observations is preferred over a design that does not.

Cornell and Gorman (1984) presented combined mixture component-process variable designs for $n \geq 3$ process variables that use only a fraction of the total number of possible design points. They considered process variables each at two levels and suggest fractions of the 2^n factorial be considered. Two plans involving the fractional factorial design in the process variables were discussed. The first plan, called a matched fraction, places the same 2^{3-1} fractional replicate design at each mixture composition point. The other plan, called a mixed fraction, uses different fractions at the composition points. Each plan was applied to the situation involving three mixture components and three process variables with the total number of design points ranging from 56 for the combined simplex-centroid by full 2^3 factorial, to only 16, which relied on running the one-quarter fraction. It should be noted that if interactions among the process variables are likely to be present, the use of a fractional factorial will result in bias being present in the coefficient estimates. Cornell and

Gorman give recommendations regarding the choice of design which depend on the form of the model to be fitted and whether or not there is prior knowledge on the magnitude of the experimental error variance.

Czitrom (1988, 1989) considered the blocking of mixture experiments consisting of three and four mixture components. She used two orthogonal blocks to construct D-optimal designs. Draper et al. (1993) consider mixture experiments with four mixture components. They treat a combination of the levels of the process variables as defining blocks. Orthogonally blocked mixture designs constructed from Latin Squares are presented. The optimal choice of a design using D-optimality is also given. While the reduction in the number of observations required can be great, obtaining D-optimality comes with a price. The D-optimal designs require very nonstandard values for the component proportions.

We propose an alternative approach to reducing the size of the experiment which borrows ideas from the above references. The concept of running only a subset of the total number of mixture-process variable combinations is borrowed from Cornell and Gorman (1984), although our fraction will involve the mixture component blends as well. To evaluate the fraction, we shall make use of the D-criterion criterion (Czitrom (1988, 1989)). The next section provides a type of experimental situation which led to this research. In the section that follows, a combined model which is slightly different in form from the combined mixture-process variable models ordinarily used is presented. The method for constructing the design and comparing it's D-criterion is discussed in the fourth section. The final section of this paper contains details on the analysis of the experiments using the proposed designs and model forms.

4.1 Experimental Situation

Historically in the mixture literature, the interest in the blending properties of the mixture components has been higher than that of studying the effects of the process variables. Generally, the process variables have been treated as "noise" factors. The primary focus on the mixture by process variable interactions has been on the effects of the process variables on the blending properties of the mixture components.

In many industrial situations, the interest in the process variables is at least equal to that in the mixture components. Consider the production of a polymer which is produced by reacting together three specific components. The research laboratory proposes a specific formulation which is the result of a highly controlled environment with reagent grade chemicals and laboratory glassware. The plant personnel use this formulation during the pilot plant and initial start-up of the full production process. During this period, the plant personnel are trying to find the proper processing conditions to produce a useable product profitably.

At some point, plant personnel need to reevaluate the polymer's formulation in light of the actual raw materials and the plant's full scale production capabilities. Plant personnel need to find the "optimal" combination of the formulation and processing conditions.

Traditionally, in response surface applications, the model assumed for process optimization is a second-order Taylor series. Such an assumption is based on background knowledge in knowing the true surface over the experimental region can be approximated by fitting a second-order model. Furthermore, in our polymer example, all second-order terms involving mixture components, process variables, and the mixture

by process variable interactions are of equal importance. In fact, the specific mixture component by process variable interaction terms may provide a significant amount of insight into which operating conditions are optimal. For instance, the engineer truly needs to know if a specific mixture component makes the reaction especially sensitive to the reaction temperature.

This type of experimental situation leads us to propose a new model for extracting information from a mixture experiment with process variables. The time and cost constraints faced by plant personnel leads us to propose a new class of designs based upon this model.

4.2 The Combined Mixture Component-Process Variable Model

In mixture experiments involving process variables, the form of the combined model consisting of terms in the mixture proportions as well as in the process variables depends on the blending properties of the mixture components, the effects of the process variables, and any interactions between the mixture components and process variables. These models are typically second-order models that allow for pure quadratic and two-factor interaction terms.

The general second-order polynomial in q mixture components is

$$\eta = \beta_0 + \sum_{i=1}^{q} \beta_i x_i + \sum_{i=1}^{q} \beta_{ii} x_i^2 + \sum\sum_{i<j}^{q} \beta_{ij} x_i x_j \,. \tag{4}$$

Now using the constraints

$$\sum_{i=1}^{q} x_i = 1 \quad \text{and} \quad x_i^2 = x_i \left(1 - \sum_{\substack{j=1 \\ j \neq i}}^{q} x_j \right),$$

Equation (4) becomes

$$\begin{aligned}\eta &= \beta_0\left(\sum_{i=1}^{q} x_i\right) + \sum_{i=1}^{q}\beta_i x_i + \sum_{i=1}^{q}\beta_{ii}x_i\left(1 - \sum_{j\neq i}^{q} x_j\right) + \sum\sum_{i<j}^{q}\beta_{ij}x_i x_j \\ &= \sum_{i=1}^{q}\left(\beta_0 + \beta_i + \beta_{ii}\right)x_i - \sum_{i=1}^{q}\beta_{ii}x_i\sum_{j\neq i}^{q} x_j + \sum\sum_{i<j}^{q}\beta_{ij}x_i x_j \\ &= \sum_{i=1}^{q}\beta_i^* x_i + \sum\sum_{i<j}^{q}\beta_{ij}^* x_i x_j\,, \end{aligned} \tag{5}$$

where $\beta_i^* = \beta_0 + \beta_i + \beta_{ii}$ and $\beta_{ij}^* = \beta_{ij} - \beta_{ii} - \beta_{jj}$ for $i, j = 1, 2, \ldots, q, i < j$.

Suppose that an experiment is to be performed with q mixture components, $x_1, x_2, \ldots, x_q$, and n process variables, $z_1, z_2, \ldots, z_n$. In the process variables, let us consider the model

$$\eta_{PV} = \alpha_0 + \sum_{k=1}^{n}\alpha_k z_k + \sum\sum_{k<l}^{n}\alpha_{kl}z_k z_l \tag{6}$$

Then there are two main types of combined models (see Cornell (1990)) that can be used in this situation. The first type is a model which crosses the mixture model terms in Equation (5) with each and every term of Equation (6). This produces the combined model

$$\begin{aligned}\eta(\mathbf{x},\mathbf{z}) &= \sum_{i=1}^{q}\beta_i^* x_i + \sum\sum_{i<j}^{q}\beta_{ij}^* x_i x_j + \sum_{i=1}^{q}\sum_{k=1}^{n}\gamma_{ik}x_i z_k + \sum_{i=1}^{q}\sum\sum_{k<l}^{n}\gamma_{ikl}x_i z_k z_l \\ &+ \sum\sum_{i<j}^{q}\sum_{k=1}^{n}\gamma_{ijk}x_i x_j z_k + \sum\sum_{i<j}^{q}\sum\sum_{k<l}^{n}\gamma_{ijkl}x_i x_j z_k z_l \end{aligned} \tag{7}$$

which includes parameters for three and four factor interactions. Depending on the design, the model of Equation (7) provides a measure of the linear and nonlinear blending properties of the mixture components averaged across the settings of the process variables as well as the effects of the process variables on the linear and nonlinear blending properties.

The second type of combined model is the additive model which combines the models in Equations (5) and (6) without crossing any of the x_i and z_j terms. This produces the model

$$\begin{aligned}\eta(\mathbf{x},\mathbf{z}) &= \sum_{i=1}^{q}\beta_i^* x_i + \sum\sum_{i<j}^{q}\beta_{ij}^* x_i x_j \\ &+ \sum_{k=1}^{n}\alpha_k z_k + \sum\sum_{k<l}^{n}\alpha_{kl} z_k z_l \,. \end{aligned} \tag{8}$$

Equation(8) provides a measure of the quadratic blending of the mixture components on the response as well as up to two-factor interactions between the process variables on the response. Since the model does not contain any crossproduct terms between the mixture components and the process variables, when fitting Equation (8) the user assumes the blending of the mixture components is the same at all factor-level combinations of the process variables. This assumption is probably unrealistic in most situations. Also, in some experiments like the one described in the previous section, the mixture component by process variable interactions may be the most important terms in the model.

A major concern with mixture experiments involving process variables is their size. Many industrial situations require the use of small experiments due to time and/or cost constraints. As the number of mixture components and/or process variables increases, the model in Equation (7) will require a design with a large number of points. While the fitting of the model in Equation (8) permits the use of a smaller design than the fitting of the model in Equation (7), it does not, as pointed out earlier, address the estimation of the mixture components by process variable interactions. If cost constraints limit the size of the experiment yet interactions between

mixture components and process variables are believed to be important, some sort of compromise between these two models is needed.

Most of the model forms that have been proposed for response surface investigations are based on a Taylor series approximation. In keeping with this tradition, suppose that the true model for the n process variables is a second-order model. Instead of Equation (6), such a model would be

$$\eta_{PV} = \alpha_0 + \sum_{k=1}^{n} \alpha_k z_k + \sum_{k=1}^{n} \alpha_{kk} z_k^2 + \sum\sum_{k<l}^{n} \alpha_{kl} z_k z_l \,. \tag{9}$$

Equation (9) is Equation (6) plus the n pure quadratic terms. Also, a Taylor series approximation for a combined second-order model would include only up to two factor interactions and would not be the model in Equation (7). Combining Equation (5) with Equation (9), our proposed combined second-order model is

$$\begin{aligned} \eta(\mathbf{x}, \mathbf{z}) \quad = \quad & \sum_{i=1}^{q} \beta_i^* x_i + \sum\sum_{i<j}^{q} \beta_{ij}^* x_i x_j + \sum_{k=1}^{n} \alpha_{kk} z_k^2 \\ + \quad & \sum\sum_{k<l}^{n} \alpha_{kl} z_k z_l + \sum_{i=1}^{q} \sum_{k=1}^{n} \gamma_{ik} x_i z_k \end{aligned} \tag{10}$$

which includes the mixture model, plus pure quadratic as well as two-factor interaction effects among the process variables, and two-factor interactions between the linear blending terms in the mixture components and the main effect terms in the process variables. The minimum number of design points needed for the proposed model (10) is less than what is needed for the completely crossed model (7) but is more than is needed for the additive model (8). Also, the proposed model can be used even if one does not feel the need for pure quadratic terms in the process variables by simply omitting those n terms.

To support the fitting of Equation (10) we shall require a design that will support nonlinear blending of the mixture components as well as the fitting of the full second order model in the process variables. In the next section, we discuss a design approach that will accomodate these terms.

4.3 Design Approach

In mixture experiments as in most response surface investigations, the design and the form of the model to be fitted go hand in hand. For example, if a second order model is suspected, it is necessary to select a design that will support the fit of this model. The design chosen must have at least as many points as there are parameters in the model. Therefore, a $(q+n)(q+n+1)/2$ point design is needed to support the fitting of the model in Equation (10).

A popular response surface design for fitting a second-order model of the form in Equation (9) is the central composite design (ccd) which consists of a complete 2^n (or a Resolution V fraction of a 2^n) factorial design, $2n$ axial points with levels $\pm\alpha$ for one factor and zero for the rest, and at least one center point. If $\alpha = 1$ is selected, the design region is a hypercube.

The approach to reducing the number of observations needed in a mixture experiment begins with a ccd in the process variables. A simplex is then placed at each point in the ccd with only a fraction of the mixture blends at each point. The mixture blends at each design point are selected from the full simplex-centroid. A general notion of balance among the mixture components across the process variables is desired. First of all, let us insist on the same number of mixture blends to be present

at both the high and low levels of each process variable. Secondly, let us insist on all of the mixture blends be present at each ± 1 factorial level for each process variable. These ideas seem very intuitive and lead us to select some of the mixture blends to be used at certain design points and different mixture blends to be used at other design points.

Two designs are considered for the fitting of the model in Equation (10). With both designs, the vertices of the simplex are run at one-half of the 2^n factorial points in the process variables with the midedge points of the simplex being run at the other half. This is done in a such a way, that if the design is collapsed across the levels of each process variable then one gets a simplex with vertices and midpoints at both the low and high level of the remaining process variables. Hence, the information in the mixture blends is spread evenly among the process variables. This is intuitively appealing since if a process variable is deemed negligible then there is still complete information on the mixture blends for the other process variables. Next, the axial points in the process variables are paired with just the centroid of the simplex. This allows for the centroid to also be present if the design is collapsed. The two designs differ only in the number of points placed at the center of the process variables. With one design the entire simplex-centroid is performed at the center while with the other only the centroid mixture blend is performed at the center of the process variables.

Consider an example involving three mixture components and two process variables. The model for this example, using Equation (10), contains 15 terms. The two designs are shown in Figures 2 and 3. For three mixture components, the design with the full simplex-centroid at the center of the process variables consists of 23 points

while the second design with just the centroid consists of 17 points. Either could be used to estimate the 15 terms in the model.

The designs in Figures 2 and 3 can be extended to experiments involving more than 3 mixture components (MC) and/or more than 2 process variables (PV). The extension is straightforward. Following the same general notion of balance described earlier, one can generate the needed designs. In this paper, a total of five cases are discussed: 3 MC, 2 PV; 3 MC, 3 PV; 4 MC, 2 PV; 4 MC, 3 PV; and 3 MC, 2 PV with upper and lower bound constraints on the mixture component proportions. For four mixture components, there are four vertices and six edge midpoints of the tetrahedron. For three process variables, the layout is a cube with $2^3 = 8$ factorial points, six axial points, and a center point. Placing upper and lower bound constraints on the mixture component proportions creates a more complicated mixture region than the simplex. The constrained region is typically an irregular polygon. The example in this paper (3 MC, 2 PV) uses the following constraints:

$$0.25 \leq x_1 \leq 0.40 \qquad 0.25 \leq x_2 \leq 0.40 \qquad 0.25 \leq x_3 \leq 0.40\,.$$

The resulting mixture region is a hexagon. Generally, the original components are transformed to L-pseudocomponents, $x_i' = (x_i - L_i)/(1 - \sum_{i=1}^{q} L_i)$ $i = 1, 2, \ldots, q$, to make the construction of the design and the fitting of the model easier. For the example in this paper, the mixture components can be transformed to L-pseudocomponents using

$$x_i' = \frac{x_i - 0.25}{1 - (0.25 + 0.25 + 0.25)} = \frac{x_i - 0.25}{0.25}, \quad i = 1, 2, 3.$$

Candidate points for the two designs in this case consist of the six vertices and

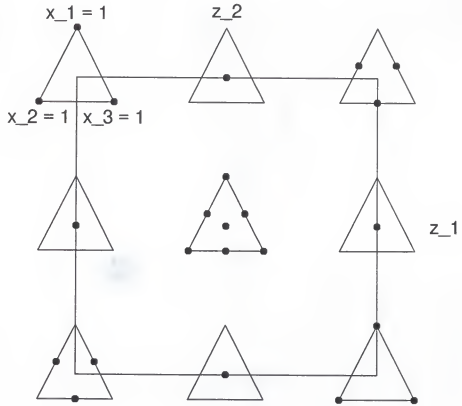


Figure 2: Proposed Design for the 3-2 Case With Full Simplex

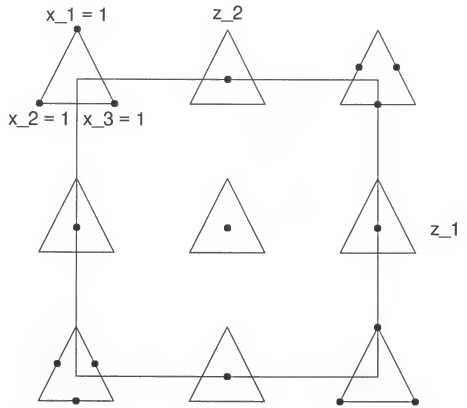


Figure 3: Proposed Design for the 3-2 Case With Just the Centroid

midpoints of the six edges of the hexagon plus the centroid at each of the nine design points of the process variables. The designs in the L-pseudocomponents are given in Appendix B.

The five cases are not intended to be exhaustive, but rather are used because they encompass a typical industrial experiment. Also, they can be used to evaluate the performance of the designs and to discuss their analysis. It is assumed that the designs can be extended to higher dimensions without complications although they may be difficult to view geometrically. To help illustrate the extension of the designs, the simplex-centroid-ccd design for the 3 MC, 3 PV case is shown in Figure 4. Each ccd axial point contains the centroid of the simplex while the ccd center point contains either the 7-point simplex-centroid or just the centroid. The coordinates of the two proposed designs for all five cases are listed in Appendix A.

When the number of design points is limited to being less than the total number generated by crossing the ccd in the process variables with the full design in the mixture components then the typical user will rely on the computer to generate a design based on some optimality criterion, such as D-optimality. Such generated designs, while optimal statistically speaking, may not be very intuitively appealing. There could be a design that is close to optimal, but has some other nice properties such as symmetry and near orthogonality. In other words, the statistician should not use the computer generated design blindly. Snee (1985) discusses some practical aspects of choosing computer-aided designs. The main point he makes is that while computer-aided designs should be used with caution, they can be helpful when the

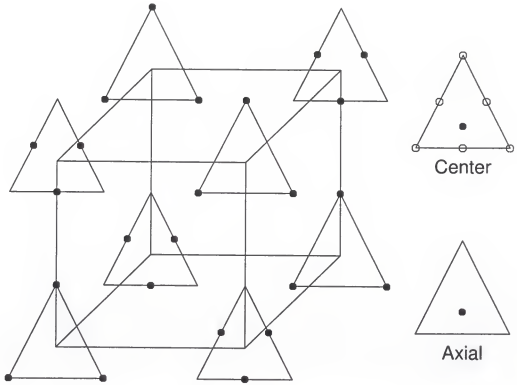


Figure 4: Proposed Design for the 3-3 Case

experimental region is irregular or when the number of available runs is limited or for special models such as those that contain nonlinear parameters.

Computer generated designs are typically not unique. Therefore, if the computer is used to generate two or three designs with the same number of points, the designs will contain some points that are not common to all the designs. Because of this, one design may be more appealing than the others for one reason or another. This leads to the question: "Wouldn't it be better to generate an acceptable or appealing design and compare it to the computer generated design?" This is what is done in this paper. Another idea is to use the computer to generate a design based on some criterion and then alter it slightly so that it is more appealing, and finally compare it back to the original design. Ultimately, if the computer produces a design that is unacceptable, an acceptable design will have to be constructed anyway. The computer can be used to provide an initial design as a start toward generating a good design. The basic idea is to use the computer generated designs wisely.

The designs presented in this paper are compared to the designs chosen by PROC OPTEX (1989) in SAS. PROC OPTEX requires the user to provide a candidate point list and the model to be fit. The candidate point list for the examples in this paper consists of the simplex-centroid design at each point of the ccd. With the 7-point simplex-centroid design in three components and a 9-point ccd in two process variables, for example, the candidate point list for three mixture components and two process variables consists of $7 \times 9 = 63$ points. The model to be fit is the model in Equation (10) with 15 terms. PROC OPTEX uses a random seed and performs ten searches from the random starting points. A few options are used. The first is to

specify the number of points in the final design, N. Then the number of points in the final design chosen by the computer is fixed at N or the default, which is the number of terms in the model, $p = 15$, plus ten additional points. A second option is the DETMAX option. This option uses the DETMAX routine developed by Mitchell (1974) to obtain a design with the maximum determinant of $\mathbf{X}'\mathbf{X}$, where $\mathbf{X}$ is a $N \times p$ matrix containing the design runs. PROC OPTEX returns the chosen design and it's D-criterion, which is defined as

$$\text{D-criterion} = 100 \times \left[\frac{\det (\mathbf{X}'\mathbf{X})^{1/p}}{N} \right]$$

and is used to compare the designs.

Two designs generated by PROC OPTEX are shown in Figures 5 and 6. These designs consist of 23 and 17 points, respectively, and are to be compared to our designs in Figures 2 and 3, respectively. Note that unless PROC OPTEX is run with the same options and starting seed, the designs generated could be different. Tables 35 and 36 show the D-criterion for our proposed designs and the designs generated by PROC OPTEX for the cases where $q = 3$ or 4 and $n = 2$ or 3. The designs from PROC OPTEX are with N restricted to the number of points from the proposed design as well as the default N. All five cases under consideration are given.

The relative efficiencies of our two proposed designs are defined as

$$\frac{\text{D-criterion for the proposed design}}{\text{D-criterion for PROC OPTEX designs with same number of points}},$$

and are given in Table 37. From Table 37, it is seen that in all cases, both of our designs are at least 73% as efficient as the computer generated designs and excluding

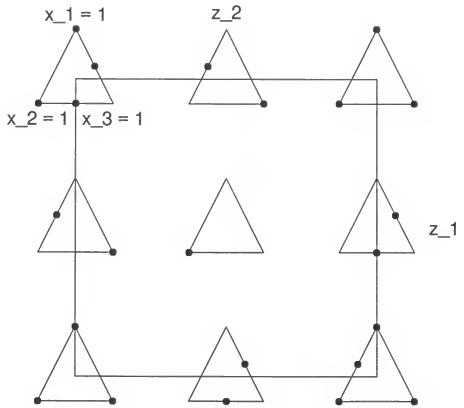


Figure 5: 23 Point Design Generated by SAS for the 3-2 Case

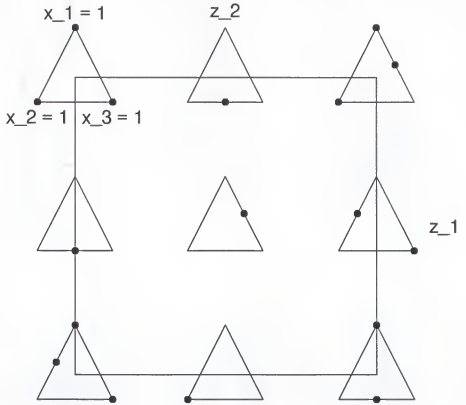


Figure 6: 17 Point Design Generated by SAS for the 3-2 Case

Table 35: Comparison of Our Design With a Full Simplex at the Center to the Designs Chosen by PROC OPTEX

Mixture Components	Process Variables	D-criterion PROC OPTEX (N-restricted)	PROC OPTEX (N = default)	Our Design (N)
3	2	11.21 (23)	11.18 (25)	8.24 (23)
3	3	14.62 (37)	14.21 (31)	12.33 (37)
4	2	5.91 (35)	5.79 (31)	4.66 (35)
4	3	8.32 (57)	7.94 (38)	7.01 (57)
Upper and Lower Constraint				
3	2	5.12 (41)	5.01 (25)	4.25 (41)

Table 36: Comparison of Our Design With Just the Centroid at the Center to the Designs Chosen by PROC OPTEX

Mixture Components	Process Variables	D-criterion PROC OPTEX (N-restricted)	PROC OPTEX (N = default)	Our Design (N)
3	2	10.26 (17)	11.18 (25)	7.70 (17)
3	3	14.21 (31)	14.21 (31)	13.11 (31)
4	2	5.39 (25)	5.79 (31)	4.55 (25)
4	3	8.14 (47)	7.94 (38)	7.47 (47)
Upper and Lower Constraint				
3	2	5.07 (29)	5.01 (25)	4.22 (29)

the 3-MC, 2-PV case, the efficiencies are even higher. Also, if one simply turned the computer loose to choose a design, the design efficiencies for almost all of our designs increases. The design with just the centroid blend at the center setting of the process variables is more efficient than when running the simplex-centroid design at the center setting of the process variables, especially in the cases involving three process variables. Hence, the addition of the mixture vertices and mid-edge points at the center of the process variables does not seem to be beneficial in terms of D-criterion. However, when considering the analysis, it is appealing to run the entire simplex-centroid at some location. Therefore, as in most design situations, there are trade-offs when deciding between designs. A plot of the D-criterion versus the size of the design is shown in Figure 7. As the design size increases, the D-criterion increases in value up to a certain point and then levels off. In other words, there is a point of diminishing return when considering the model in Equation (10).

Suppose the model in Equation (6) is used for the process variables instead of the model in Equation (9) (i.e. the model for the process variables is believed to be first-order plus interactions). As mentioned earlier, the combined model can be adjusted by omitting the n pure quadratic terms in the process variables. Then the number of points in the two proposed designs can easily be reduced for this situation. All that needs to be done is to eliminate the axial points which are no longer required since the pure quadratic terms in the z_j's are no longer in the model. This results in a reduction in design size of $2n$ points for n process variables. This flexibility is another appealing feature of the two proposed designs.

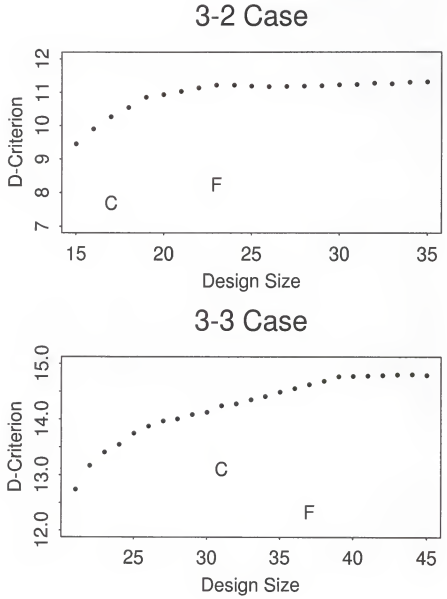


Figure 7: Plot of D-Criterion Versus Design Size (C: Design With Just the Centroid, F: Design With the Full Simplex-Centroid at Center of Process Variables)

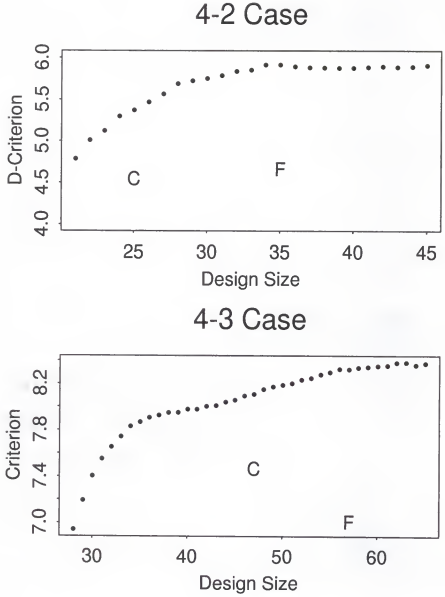


Figure 7: Continued

Table 37: Relative Efficiencies of the Proposed Designs As Compared to the Designs With the Same Number of Points Chosen by PROC OPTEX

Mixture Components	Process Variables	Relative Efficiency: Full Simplex	Relative Efficiency: Just Centroid
3	2	73.5 †	75.1 ††
3	3	84.3 †	92.3
4	2	78.8 †	84.4 ††
4	3	84.3 †	91.8 †
Upper and Lower Constraint			
3	2	83.0 †	83.2 †

† If relative efficiency is defined using the default N in the denominator, relative efficiencies would increase.

†† If relative efficiency is defined using the default N in the denominator, relative efficiencies would decrease.

4.4 Analysis

Cornell (1990) discusses the analysis for the additive and crossed models that use standard designs. In this paper, the appropriate analysis upon fitting the compromised model given in Equation (10) using either of the two proposed designs is needed. It will be assumed throughout this section that the design is run as a completely randomized design.

For the proposed designs, there is a total of $N-1$ degrees of freedom. Estimating the terms in the model requires $p-1$ df. Using effect sparsity, the remaining $N-p$ df can be pooled to form an error source which can be used to test the significance of the terms in the model. The ANOVA table is shown in Table 38.

Table 38: ANOVA Table for the Proposed Designs Using the Compromised Model

Source		DF
mixture components:		
linear	$q-1$	
quadratic	$q(q-1)/2$	
process variables:		
2-factor int.	$n(n-1)/2$	
quadratic	n	
MC × PV int.:		
2-factor	$q \cdot n$	
Total Effects		$p-1=\frac{(q+2)(q-1)+n(n+1)+2qn}{2}$
Error		$N-p$
Total		$N-1$

Table 39: Error Degrees of Freedom for the Two Proposed Designs Under the 5 Cases

		Error DF	
Mixture Components	Process Variables	Full Simplex	Just Centroid
3	2	8	2†
3	3	16	10
4	2	14	4†
4	3	29	19
Upper and Lower Constraint			
3	2	26	14

† The df may be insufficient for testing. It is recommended that either the design with the full simplex be used or replicates of the centroid be taken at the center of the process variables.

In most cases, both of the proposed designs have sufficient df to estimate the model. However, for testing the significance of the terms in the model, the design using the full simplex-centroid at the center of the process variables has more df for error. Table 39 lists the model and error df for the two proposed designs for the five cases under consideration. For the unconstrained 3-MC, 2-PV and the 4-MC, 2-PV cases, there may be insufficient df for error for the design with the single centroid point, and tests of significance in these situations are not very powerful. In these cases, it is recommended that either the design with the simplex-centroid at the center of the process variables be used or replicates of the centroid blend be taken at the center of the process variables to estimate the error variance.

The testing of the effects should begin with the mixture component by process variable interactions. These interactions represent the effects of the process variables

on the linear blending properties of the mixture components. Once these have been investigated, the nonlinear blending properties of the mixture components and the quadratic effects of the process variables can be tested and this may result in changing the terms in the original model. The investigation of the interaction terms involves a main question and the two subsequent questions:

- Main: Is the effect of the process variable the same for all blends of the mixture components?

    1. If so, is this effect significantly different from 0?
    2. If not, where and how is the effect different?

for each of the n process variables. For a given process variable, there are q interaction terms of the form

$$\gamma_{1k}x_1z_k, \gamma_{2k}x_2z_k \cdots, \gamma_{qk}x_qz_k, \quad \text{for } k = 1, 2, \ldots, n\,.$$

The questions above partition the q df involving the q interaction terms into $q-1$ df for testing the "main" question and 1 df for testing sub-question number 1. The "main" question above asks if the effect of z_k is the same on each of the mixture components. Therefore, the hypotheses are

$$\begin{array}{ll} \mathrm{H}_0: & \gamma_{1k} = \gamma_{2k} = \cdots = \gamma_{qk} \\ \mathrm{H}_A: & \text{at least one not equal} \end{array}$$

and the tests are carried out for each process variable.

To illustrate how the tests on the interaction coefficients are carried out, consider the case of three mixture components. Then for process variable, z_1, the hypotheses

are

$$\begin{array}{ll} \mathrm{H}_0: & \gamma_{11} = \gamma_{21} = \gamma_{31} \\ \mathrm{H}_A: & \text{at least one not equal.} \end{array}$$

The testing procedure can be rewritten as the simultaneous test of

$$\begin{array}{ll} \mathrm{H}_0: & \gamma_{11} - \gamma_{21} = 0 \\ \mathrm{H}_A: & \gamma_{11} - \gamma_{21} \neq 0 \end{array} \qquad \begin{array}{ll} \mathrm{H}_0: & \gamma_{11} - \gamma_{31} = 0 \\ \mathrm{H}_A: & \gamma_{11} - \gamma_{31} \neq 0. \end{array}$$

If the three terms here are entered last in the model, then in matrix notation the above hypotheses become

$$\mathrm{H}_0 : \mathbf{L}'\boldsymbol{\beta} = \mathbf{c}$$

where $\mathbf{L}' = \left[\mathbf{0}_{2,p-3}, \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}\right]$, $\boldsymbol{\beta}$ is a $p \times 1$ vector of all the parameters in the model, and $\mathbf{c} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. The appropriate F-test is

$$F = \frac{Q/s}{\mathrm{MSE}} \overset{\mathrm{H}_0}{\sim} F_{s,N-p-1}$$

where s (=2 in our case) is the rank of $\mathbf{L}'$, MSE is the mean squared error from the fit of the model, and

$$Q = \left(\mathbf{L}'\widehat{\boldsymbol{\beta}} - \mathbf{c}\right)' \left[\mathbf{L}' \left(\mathbf{X}'\mathbf{X}\right)^{-1} \mathbf{L}\right]^{-1} \left(\mathbf{L}'\widehat{\boldsymbol{\beta}} - \mathbf{c}\right)$$

with $\widehat{\boldsymbol{\beta}}$ being the least squares estimate of the parameter vector $\boldsymbol{\beta}$. The test can easily be performed in SAS using PROC REG. After the model statement, the test statement is used with a separate test statement for each process variable. For the above example, the SAS code is

```
PROC REG;
   MODEL  Y = x1 x2 ··· x1z1 x2z1 x3z1 / noint;
   TEST   x1z1 − x2z1 = 0, x1z1 − x3z1 = 0;
   TEST   x1z2 − x2z2 = 0, x1z2 − x3z2 = 0;
   TEST   ···, ···;
   RUN;
```

would produce the correct simultaneous test of the hypotheses. If H_0 is rejected, then there is an interaction present between the mixture components and the process variable. Hence, the effect of the process variable is not the same for all mixture components. A further investigation of the form of the interaction needs to be carried out. This is easily accomplished with plots.

If H_0 is not rejected, then for a fixed process variable, k, the γ_{ik}'s are equal. In other words, the process variable affects the mixture components equally when the mixture region is a simplex. A follow-up test is needed since the equal magnitude could be zero. As it turns out, this test will investigate whether or not there is an additive effect due to the process variable.

Since $H_0 : \gamma_{1k} = \gamma_{2k} = \cdots = \gamma_{qk}$ is not rejected, the terms in the model in Equation (10) involving the interaction of the mixture components and process variable, k, can be rewritten as

$$\begin{aligned} \gamma_{1k}x_1z_k + \gamma_{2k}x_2z_k + \cdots + \gamma_{qk}x_qz_k &= \gamma_k x_1 z_k + \gamma_k x_2 z_k + \cdots + \gamma_k x_q z_k \\ &= \gamma_k z_k(x_1 + x_2 + \cdots + x_q) \\ &= \gamma_k z_k . \end{aligned}$$

Therefore, the follow-up test is a test of the significance of the additive effect of process variable k. Note, this would have been the test for the additive model. All that needs to be done for the analysis is to refit the model with z_k replacing the interactions of the mixture components with z_k. This will provide the 1 df t-test that combined with the $q - 1$ df from above partitions the q df involving the interactions of the mixture components and the process variables. This is the test for the additive effect of the process variable z_k and is done for each process variable that H_0 is not

rejected in the "main" question. If H_0 for testing z_k is not rejected, then there is no additive effect due to the process variable.

After a thorough investigation of all the mixture component by process variable interactions, the rest of the model should be studied. Even if the mixture component by process variable interactions exist, it is still beneficial to investigate the mixture component part of the model as well as the curvelinear terms involving the process variables.

To investigate the mixture component part of the model, generally the linear blending terms are not tested. However, if the nonlinear blending terms cannot be shown to be different from zero, then one can test whether the linear blending terms are all equal (see Cornell (1990), Chapter 5). Simple t-tests can be performed on the nonlinear blending coefficient estimates to ascertain which pairs of components blend nonlinearly. In other words, detecting curvature in the shape of the mixture surface through the nonlinear blending of the components is of primary interest.

When data are collected from a simplex-centroid design and curvature in the surface is detected by finding the coefficient estimate of the binary crossproduct, this generally means the binary blend produces a response value which differs from the average of the response values collected at the two vertices. If the coefficient is positive and the test is significant, this means the binary blend produces a higher response value than is obtained from simply averaging the response values from the individual components. Similarly, for a significant negative parameter estimate, the response to the binary blend is lower than the average of the responses at the vertices. The tests

for these parameters involve the least squares estimates and are performed in SAS using PROC REG.

For the process variables, in addition to testing their effects on the linear blending properties of the mixture components, tests on the pure quadratic effects and the two-factor interactions among the process variables are of interest. If the test for a pure quadratic term is significant, then the response is curvilinear as one moves from the low to the high level of that process variable. A significant interaction implies that the relationship between the two process variables is not always the same. Before eliminating any terms from the model resulting testing results being nonsignificant, the model should be tested for lack of fit.

4.5 Lack of Fit

The model in Equation (10) is assumed to be the true model. However, it is possible that this is not the case. There may be some important higher order terms that were not included in the model. Therefore, before making any inferences about significant effects, the model should be tested for lack of fit.

One way to test the model for lack of fit is to replicate one or more points in the design. Then, the error or residual sum of squares in the ANOVA table can be partitioned into sum of squares due to lack of fit and sum of squares for pure error. The ratio of these two divided by their appropriate degrees of freedom forms an F-test which can be used to determine if the model suffers from lack of fit. Another way is to use check points which involves three steps:

1. Fit the proposed model using data collected at the design points.

2. Collect additional observations at points other than the design points. These are called check points.

3. Compare the observed data at the check points to the predicted response from the fitted model at the check points.

The basic idea is that if the predictions are close to the observed values, then the model does not suffer from lack of fit (see Cornell (1990) for more details).

If lack of fit is detected, then the proposed model needs to be upgraded by the addition of higher-order terms. Significant terms in the fitted model might suggest possible candidates for lack of fit. For the proposed model, some likely candidates are

i. the special cubic terms involving the mixture components (ie. $x_1x_2x_3$ ...)

ii. the interactions between binary blends and the linear effects of a process variable (ie. $x_1x_2z_1$ ...)

iii. the interactions between the linear blending terms and the two-factor interactions among the process variables (ie. $x_1z_1z_2$ ...)

Once the necessary terms are added to the model, the new model should also be checked for lack of fit.

One nice property of our proposed designs is that both designs can support the fitting of higher order terms without adding more design points. The design with the simplex-centroid at the center of the process variables is used since it has more degrees

Table 40: Higher Order terms Supported by the Proposed Design With Three Mixture Components and the Simplex-Centroid at the Center of the Process Variables

	Binary Blend by Process Variable Interactions		
	3MC, 2PV	3MC, 3PV	Constrained 3MC, 2PV
Without Special Cubic Terms	1 Binary Blend With Both Process Variables	All 9 Terms	All 6 Terms
With Special Cubic Terms	Any 2 of the 3 Binary Blends With Either Process Variable OR 1 Binary Blend With Both Process Variables	All 9 Terms	All 6 Terms

of freedom left over for lack of fit terms than the design with just the mixture centroid at the center of the process variables. To simplify the many possible combinations of higher order terms that could be fit, only the terms in (i) and (ii), namely, the special cubic terms and the interactions between a binary blend and the linear effect of a process variable, are considered. Tables 40 and 41 give the additional terms the proposed design can accomodate for each of the five cases. It is seen from Tables 40 and 41 that the proposed design supports some of the higher order terms in most cases and all of the higher order terms in the case of three mixture components and three process variables.

If check points are going to be used to test for lack of fit, one alternative to the method described in Cornell (1990) is to add points that would aid in the fitting of higher order terms, such as face center points. If the proposed model suffers from

Table 41: Higher Order terms Supported by the Proposed Design With Four Mixture Components and the Simplex-Centroid at the Center of the Process Variables

	Binary Blend by Process Variable Interactions	
	4MC, 2PV	4MC, 3PV
Without Special Cubic Terms	The 6 Binary Blends with One of the Process Variables Or 3 Binary Blends with Both Process Variables	All 18 Terms
With Special Cubic Terms	1 Special Cubic Plus the 3 Binary Blends of the MC's Involved in the Special Cubic With Either Process Variable	All 18 Terms

lack of fit, then some of the special cubic terms and the interactions between a binary blend and a process variable can be supported by the proposed design without any additional points.

4.6 Example

Consider an experiment involving three mixture components (x_1, x_2, x_3) and two process variables (z_1, z_2). The data for this example is constructed for illustrative purposes and is given in Table 42. The analysis is carried out in the manner described in this chapter. The parameter estimates and the results of performing t-tests on the estimates are shown in Table 43.

The analysis should begin with the mixture component by process variable interactions. The "main" question needs to be answered which asks if the effect of the process variable is the same for all blends of the mixture components. The appropriate

Table 42: Data for the Example

x_1	x_2	x_3	z_1	z_2	Response
1	0	0	−1	1	4
0	1	0	−1	1	9
0	0	1	−1	1	5
1	0	0	1	−1	9
0	1	0	1	−1	15
0	0	1	1	−1	8
.5	.5	0	−1	−1	12
.5	0	.5	−1	−1	7
0	.5	.5	−1	−1	8
.5	.5	0	1	1	15
.5	0	.5	1	1	12
0	.5	.5	1	1	12
.33	.33	.33	−1	0	7
.33	.33	.33	1	0	10
.33	.33	.33	0	−1	7
.33	.33	.33	0	1	8
.33	.33	.33	0	0	9

Table 43: Original Parameter Estimates for the Example

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
X_1	1	11.43	2.74	4.172	0.0529
X_2	1	16.93	2.74	6.179	0.0252
X_3	1	11.43	2.74	4.172	0.0529
X_1X_2	1	−6.45	6.45	−1.000	0.4228
X_1X_3	1	−11.45	6.45	-1.775	0.2179
X_2X_3	1	−20.45	6.45	−3.171	0.0867
Z_1Z_1	1	−0.50	1.07	−0.468	0.6858
Z_2Z_2	1	−1.50	1.07	−1.404	0.2955
Z_1Z_2	1	2.93	0.79	3.681	0.0665
X_1Z_1	1	2.15	0.61	3.536	0.0715
X_1Z_2	1	−0.15	0.61	−0.254	0.8230
X_2Z_1	1	1.90	0.61	3.125	0.0889
X_2Z_2	1	−0.90	0.61	−1.485	0.2759
X_3Z_1	1	2.15	0.61	3.536	0.0715
X_3Z_2	1	0.85	0.61	1.387	0.2998

test statement is written in SAS and produces

F value: 0.0438 Prob>F: 0.9581

for process variable one and

F value: 1.6213 Prob>F: 0.5056 .

Therefore, none of the interactions between the mixture components and the process variables are significant. The follow-up question asks if the process variable effects are different from zero or additive effects. For process variable one, the p-value is 0.0012 and for process variable two it is 0.8611. This leads to an additive effect for process variable one. It turns out that after eliminating the terms involving z_2, there is also an additive effect for z_1^2. Finally, there is some evidence of positive nonlinear blending of x_1 and x_2.

4.7 Summary

In this chapter, we have introdced a new class of combined designs for mixture experiments with process variables. These designs are specifically created to support a second-order Taylor series approximation of the true response function. These designs employ an intuitive balance which is very appealing and also support the fitting of some lack of fit terms. On the whole, they are quite competitive to D-optimal designs generated by common software.

In addition, the combined second-order model proposed in terms of both the mixture components and the process variables represents an explicit compromise between the mixture and standard response surface schools of thought. The proposed

model allows the experimenter to see directly the interaction between specific mixture components and specific process variables. These interactions can provide valuable insights into the entire process being studied.

CHAPTER 5
MIXTURE EXPERIMENTS WITH PROCESS VARIABLES IN A SPLIT-PLOT SETTING

With many industrial experiments, it is often difficult or costly to run the experiment in a completely random order. Often times, a level of one or more factors is fixed and then all or a fraction of all of the combinations of the other factors are run. This process is repeated until the desired number of runs have been used. The result is a design that uses restricted randomization and resembles a split-plot design.

Mixture experiments with process variables are often run under the above setting. Typically, the process variables serve as the whole plot factors and the mixture components make up the subplot factors. However, in some experiments their roles could be reversed. For a discussion of mixture experiments run using split-plot designs, see Cornell (1988).

Consider the proposed model and class of designs for a mixture experiment with process variables presented earlier. In this chapter, we extend the ideas presented in Chapter 4 to the situation where complete randomization is not possible. This may be due to physical constraints of the experiment or to cost constraints. We will assume that the experiment is conducted by fixing the levels of the process variables and then running some of the mixture blends. Therefore, the process variables are the whole plot factors while the mixture components are the subplot factors. The goal of this chapter is to consider different methods of estimation for the effects under the restricted randomization.

5.1 First-Order Model in the Process Variables

First, we shall consider the case where a first-order plus ionteractions model is assumed for the process variables which are to serve as the whole plot factors. Therefore, the model for the process variables is the model given in Equation (6) of Chapter 4. When this model is combined with the model in Equation (5) of Chapter 4 under a Taylor series approach, the resulting model is

$$\eta(\mathbf{x}, \mathbf{z}) = \sum_{i=1}^{q} \beta_i^* x_i + \sum\sum_{i<j}^{q} \beta_{ij}^* x_i x_j + \sum\sum_{k<l}^{n} \alpha_{kl} z_k z_l + \sum_{i=1}^{q}\sum_{k=1}^{n} \gamma_{ik} x_i z_k \,.$$

Now, a design is needed that will support the fit of this model.

Proposed Design

The designs proposed in Chapter 4 use a ccd in the process variables with a fraction of the mixture blends run at each location in the ccd. The vertices of the simplex are run at half of the the 2^n factorial points in the process variables with the midedge points being run at the other half. Also, the axial points in the process variables are paired with the centroid of the simplex. Finally, either the simplex-centroid or simply the centroid is run at the center of the process variables. Figures 2 and 3 in Chapter 4 show the proposed designs for three mixture components and two process variables.

Initially in this section, we are assuming only a first order model in the process variables. Hence, it is unnecessary to include the axial points. It is important to keep in mind that the nature of the experiment is that of a split-plot experiment. This will influence how the design is chosen.

If a balanced design is desired, then the same number of subplots needs to be run in each whole plot. This means that at each of the combinations of the process variables, the same number of mixture blends should be run. The center of the process variables will be used as a replicate point in order to get estimates of the pure error for the whole plots and the subplots.

Consider the case of three mixture components and two process variables. At the $2^2 = 4$ factorial points of the process variables, either the vertices or the midedge points in the mixture components are run according to Figures 2 and 3 of Chapter 4. However, the centroid will also be included at each of the four points. This is done so that fewer replicates of the center of the process variables will be needed to get a good estimate of the pure error at the subplot level. A more thorough explanation of this will be given in the next subsection. In each whole plot there will be four subplots. At the center of the process variables, four replicates of the centroid will be included. Also, the entire whole plot at the center of the process variables will be replicated in order to get an estimate of the whole plot error variance. Figure 8 gives the proposed design.

It should be reiterated that while the design looks similar to those proposed in Chapter 4, the difference is in the way the experiment is carried out. In this chapter, the levels of the process variables are fixed and then the corresponding four mixture blends are run in some random order. Then, the levels of the process variables are fixed again and the four blends are run and so on. This of course is different from a completely randomized design and can be thought of as a split-plot experiment.

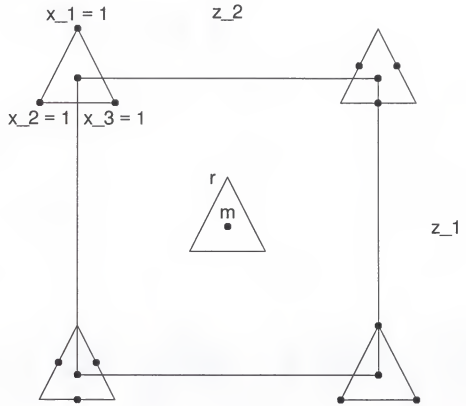


Figure 8: Proposed Design for Split-Plot Structure With a First-Order Model

The proposed design for three mixture components and two process variables can be extended to higher dimensions. For three process variables the whole plot treatments are now on the vertices of a cube. There are eight factorial points and then some replicates of the center. The case of four mixture components is not quite as straightforward since there are four vertices but six midedge points. In this case, two replicates of the centroid will be run in the whole plots that contain the vertices. Also, the center of the process variables will have six replicates of the centroid. The design will then be balanced with six subplots in each whole plot.

Estimation

Model estimation under the split-plot structure of the experiment is more complex than when the the experiment is completely randomized. The initial randomization of the process variables corresponds to the first randomization. This generates the whole plot error variance, σ_δ^2. The randomization of the mixture blends constitutes a second randomization which generates the subplot error variance, σ_ϵ^2. Hence, a model for the experiment is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\delta} + \boldsymbol{\epsilon}$$

where

$$\boldsymbol{\delta} + \boldsymbol{\epsilon} \sim N(0, \mathbf{V}) .$$

The matrix

$$\mathbf{V} = \sigma_\delta^2 \mathbf{J} + \sigma_\epsilon^2 \mathbf{I}$$

represents the variance-covariance structure of the split-plot experiment. $\mathbf{J}$ is a block

diagonal matrix of $\mathbf{1}_{b\times 1} \times \mathbf{1}'_{1\times b}$ where b is the number of observations in each whole plot and $\mathbf{I}$ is the identity matrix of dimension b. The structure implies that observations in different whole plots are independent while observations within a whole plot are correlated.

To complicate matters even more, the same mixture blends are not run at each combination of the process variables. Because of this deviation in treatment structure, ordinary least squares is not equivalent to generalized least squares. Therefore, the estimating equation is

$$\widehat{\boldsymbol{\beta}} = \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$$

with

$$\mathrm{Var}(\widehat{\boldsymbol{\beta}}) = \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1} .$$

The estimation of the model coefficients and of the variances of these estimates depends on the matrix $\mathbf{V}$ and thus on σ_δ^2 and σ_ϵ^2. Generally, the analysis requires the estimation of the two error variances. Three methods, ordinary least squares, restricted maximum likelihood and a method based on pure error, will be considered for estimating these error variances.

Ordinary least squares (OLS) assumes the observations are independent. Thus, it ignores the dependent structure of the split-plot design. This is a naive approach in light of the restricted randomization and one would expect that it will not perform well. Restricted maximum likelihood (REML) is similar to maximum likelihood estimation in that it uses the likelihood of a transformation of $\mathbf{y}$ which is based on the residuals (see Russell and Bradley (1958)). The proposed design uses r replicates of the whole plot at the center of the process variables. Within each of these whole plots,

there are m replicates of the centroid. The final method uses these replicate points to form pure error terms which can be used to estimate the two error variances.

Simulation Study

An investigation is conducted to evaluate the performances of the three methods. A simulation study is carried out to obtain approximations for the elements of the matrix, $\text{Var}(\widehat{\boldsymbol{\beta}})$. Then the det $\left[\widehat{\text{Var}}(\widehat{\boldsymbol{\beta}})\right]$ is computed and compared to the asymptotic values which use known $\mathbf{V}$. By looking at the det $\left[\widehat{\text{Var}}(\widehat{\boldsymbol{\beta}})\right]$, we are comparing the size of the joint confidence ellipsoid around the parameter estimates. The five cases from Chapter 4 are considered.

It is convenient to define the relationship

$$d = \frac{\sigma_\delta^2}{\sigma_\epsilon^2}.$$

As d increases in value greater than unity, the whole plot error variance becomes much larger than the subplot error variance and thus the correlations among the observations become stronger. Without loss of generality, we will assume that $\sigma_\epsilon^2 = 1$ and thus $d = \sigma_\delta^2$ represents the whole plot error variance. The asymptotic value for the $\text{Var}(\widehat{\boldsymbol{\beta}})$ is $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$ where $\mathbf{V} = d\mathbf{J} + \mathbf{I}$. The summary value $|(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}|$ will be used for comparing the three methods.

For OLS, the estimate of $\boldsymbol{\beta}$ is $\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$. The estimated variance of $\widehat{\boldsymbol{\beta}}$ is

$$\begin{aligned}\widehat{\text{Var}}(\widehat{\boldsymbol{\beta}}) &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\left(\text{Var}(\mathbf{y})\right)\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\end{aligned}$$

where $\mathbf{V} = d\mathbf{J} + \mathbf{I}$ as above. Therefore, the quantity $|(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}|$ can be calculated and compared to the asymptotic value.

Using REML changes the estimate of $\boldsymbol{\beta}$ to $\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{y}$ where $\mathbf{V}_c = \hat{\sigma}_\delta^2\mathbf{J} + \hat{\sigma}_\epsilon^2\mathbf{I}$. The variance is

$$\begin{aligned}\widehat{\mathrm{Var}}(\widehat{\boldsymbol{\beta}}) &= (\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}_c^{-1}\left(\mathrm{Var}(\mathbf{y})\right)\mathbf{V}_c^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{X})^{-1} \\ &= (\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{V}\mathbf{V}_c^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{X})^{-1}\end{aligned}$$

and the determinant is used for comparison. Therefore, better estimates of σ_δ^2 and σ_ϵ^2 will make the determinant of the REML estimate of the variance of $\widehat{\boldsymbol{\beta}}$ closer to the determinant of the asymptotic value. Clearly, in practice $\mathbf{V}$ is not available to the experimenter. However, we do know $\mathbf{V}$ for our simulation. Hence, we can calculate the correct variance-covariance structure.

To obtain $\mathbf{V}_c$, a simulation is conducted under the assumption that $\sigma_\epsilon^2 = 1$. Then, for a fixed known value of d, $\mathbf{V}_c$ can be computed. Since there is no prior knowledge of which terms in the model will be important, all model coefficients should be assumed equal. For convenience, they will be set equal to 1 (ie. $\boldsymbol{\beta} = \mathbf{1}$). The simulation of the values needed to compute $\mathbf{V}_c$ is a two-step procedure corresponding to the split-plot nature of the experiment. The whole plot is simulated by

$$W = f(\mathbf{z}) + N(0, d)$$

where $f(\mathbf{z})$ consists of the whole plot terms in the model evaluated at the point $\mathbf{z}$ and $N(0, d)$ is the whole plot error term. Then, the observations are given by

$$\mathbf{y} = f(\mathbf{x}; \mathbf{z}) + W + N(0, 1)$$

where $f(\mathbf{x}; \mathbf{z})$ consists of the subplot terms and the whole plot $\times$ subplot interactions in the model evaluated at the point $(\mathbf{x}; \mathbf{z})$ and $N(0, 1)$ is the subplot error term.

This is done for each whole plot. These observations along with the design matrix and the model are inputted into PROC MIXED in SAS. Using the code provided in Letsinger, Myers and Lentner (1996) and given in Appendix C, the estimated variance components, $\hat{\sigma}_\delta^2$ and $\hat{\sigma}_\epsilon^2$, are computed. These estimates are used to get $\mathbf{V}_c$. Finally, the determinant

$$\left|(\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{V}\mathbf{V}_c^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}_c^{-1}\mathbf{X})^{-1}\right|$$

is calculated. This procedure is repeated 10,000 times and the average determinant is used for comparison to the asymptotic determinant.

For the method using pure error estimates, simulation is also used. However, the only response values that need to be simulated are at the replicate points. The simulation is carried out the same way as described for the REML case. In other words, first $W = f(\mathbf{z}) + N(0, d)$ is simulated and then $\mathbf{y} = f(\mathbf{x}; \mathbf{z}) + W + N(0, 1)$ is obtained for each whole plot.

A total of r replicates at the center of the process variables are used. Since a balanced design is desired, in each of these whole plots a total of m replicates of the centroid are used. This leads to the $rm - 1$ total degrees of freedom being partitioned into $r - 1$ df for the pure error term involving the process variables (whole plot factors) and $r(m-1)$ df for the pure error term involving the mixture components (subplot factors). The estimated whole plot pure error term is

$$S_{wp}^2 = \frac{\sum_{i=1}^{r}\left(\overline{Y}_{i.} - \overline{Y}_{..}\right)^2}{r-1}$$

where $\overline{Y}_{i.}$ is the mean of the simulated values in the i^{th} whole plot and $\overline{Y}_{..}$ is the overall mean of the simulated values. From the replicated centroids at the center of

the process variables, the estimated subplot pure error term is

$$S^2_{sp} = \frac{\frac{1}{r}\sum_{i=1}^{r}\sum_{j=1}^{m}\left(Y_{ij} - \bar{Y}_{i.}\right)^2}{m-1}.$$

When there are three mixture components, the above term is the estimated subplot error term. However, when there are four mixture components, two total replicates of the centroid are added to the whole plots that contain the vertices. As was pointed out earlier, this is done to have a balanced design. So in each of these whole plots, one extra degree of freedom can be used to estimate the subplot error term. The estimate of the subplot error term becomes a weighted error term and is given by

$$S^2_{sp} = \frac{(2-1)S^2_{sp,V1} + (2-1)S^2_{sp,V2} + \cdots + (2-1)S^2_{sp,V2^{k-1}} + S^2_{sp,0}}{(2-1) + (2-1) + \cdots + (2-1) + r(m-1)}$$

$$S^2_{sp} = \frac{S^2_{sp,V1} + S^2_{sp,V2} + \cdots + S^2_{sp,V2^{k-1}} + S^2_{sp,0}}{2^{k-1} + r(m-1)}$$

where

$$S^2_{sp,0} = (m-1)\left(\frac{\sum_{i=1}^{r}\sum_{j=1}^{m}\left(Y_{ij} - \bar{Y}_{i.}\right)^2}{m-1}\right)$$

and $S^2_{sp,Vk} = (Y_1 - \bar{Y})^2 + (Y_1 - \bar{Y})^2$ is the estimate of subplot error from the two replicated centroids in the whole plots that have the vertices, $k = 1, 2, \ldots, 2^{k-1}$.

The variance of $\widehat{\boldsymbol{\beta}}$ is the same as given for the REML case. Once the pure error estimates are obtained using the above formulas, they can be used to get $\mathbf{V}_c$. Again, the procedure is repeated 10,000 times and the $\det(\widehat{\text{Var}}(\widehat{\boldsymbol{\beta}}))$ is calculated each time. Then, the average of the determinants is compared to the determinant of the asymptotic $\text{Var}(\widehat{\boldsymbol{\beta}})$.

Comparison of Methods

The methods are compared for various values of d. Since it is assumed that $\sigma_\epsilon^2 = 1$, then $d = \sigma_\delta^2$ represents the whole plot error variance as well as the ratio of the two error terms. Following Letsinger, Myers and Lentner (1996), five values are used for d: .11, .43, 1.0, 2.3, and 4.0. This includes values where the whole plot error variance is smaller than (.11, .43), equal to (1.0) and greater than (2.3, 4.0) the subplot error variance. The designs consisted of $r = 2$, 3, 4 and 5 total replicates of the center of the process variables (whole plot factors).

Tables 44 - 48 show the asymptotic value, the OLS value, the average determinant values for REML, and the average determinant values for the pure error approach. The standard errors for the simulated values are also given. It appears that in general the standard errors are smaller for the pure error method especially for 4 and 5 replicates of the center of the process variables. Since it is difficult to compare the average determinants directly, the notion of relative efficiency is used.

The results from the simulation are compared to the asymptotic values. This is done through relative efficiency which for OLS is defined as

$$\text{Rel. Eff.} = \frac{|(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}|}{|(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}|}$$

where $\mathbf{V}$ is the true variance-covariance matrix. For REML and the pure error method, relative efficiency is defined as

$$\text{Rel. Eff.} = \frac{\text{avg.}\,|(\mathbf{X}'\mathbf{V}_e^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}_e^{-1}\mathbf{V}\mathbf{V}_e^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}_e^{-1}\mathbf{X})^{-1}|}{|(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}|}$$

where the average is over the 10,000 simulated determinants. Tables 49 - 53 show the results for the five values of d and $r = 2$, 3, 4, and 5.

Table 44: Values for the Variance of $\widehat{\boldsymbol{\beta}}$ (3 Mixture Components and 2 Process Variables)

d	Asymptotic	OLS	REML	Pure Error
		$r = 2$		
.11	0.01	0.01	0.01(.00)	0.01(.00)
.43	0.20	0.25	0.22(.00)	0.22(.00)
1.0	2.54	4.51	3.18(.00)	3.04(.00)
2.3	46.98	142.17	64.98(.31)	65.89(.29)
4.0	372.51	1742.28	553.95(3.94)	606.82(3.96)
		$r = 3$		
.11	0.01	0.01	0.01(.00)	0.01(.00)
.43	0.16	0.20	0.18(.00)	0.01(.00)
1.0	2.09	3.72	2.53(.00)	2.33(.00)
2.3	39.51	121.40	48.90(.21)	46.91(.14)
4.0	315.56	1502.61	420.35(2.75)	393.39(1.96)
		$r = 4$		
.11	0.01	0.01	0.01(.00)	0.01(.00)
.43	0.13	0.17	0.15(.00)	0.14(.00)
1.0	1.78	3.21	2.02(.00)	1.92(.00)
2.3	34.08	104.63	40.24(.13)	38.25(.08)
4.0	273.71	1301.68	313.64(1.32)	311.67(.99)
		$r = 5$		
.11	0.01	0.01	0.00(.00)	0.01(.00)
.43	0.11	0.14	0.12(.00)	0.12(.00)
1.0	1.55	2.79	1.76(.00)	1.65(.15)
2.3	29.97	91.52	33.35(.08)	32.40(.05)
4.0	241.66	1142.11	277.25(1.11)	263.17(.64)
Standard Error for the Simulation are in Parentheses				

Table 45: Values for the Variance of $\hat{\boldsymbol{\beta}}$ (3 Mixture Components and 3 Process Variables)

d	Asymptotic	OLS	REML	Pure Error
			$r = 2$	
.11	7.1×10^{-12}	7.3×10^{-12}	$8.6 \times 10^{-12}(1.3 \times 10^{-14})$	$7.5 \times 10^{-12}(7.3 \times 10^{-15})$
.43	7.7×10^{-10}	9.2×10^{-10}	$8.7 \times 10^{-10}(7.9 \times 10^{-13})$	$8.7 \times 10^{-10}(7.5 \times 10^{-13})$
1.0	6.4×10^{-8}	1.2×10^{-7}	7.2×10^{-8} (1.3×10^{-10})	8.2×10^{-8} (1.8×10^{-10})
2.3	1.1×10^{-5}	3.2×10^{-5}	1.3×10^{-5} (5.3×10^{-8})	1.7×10^{-5} (7.4×10^{-8})
4.0	3.9×10^{-4}	1.9×10^{-3}	4.9×10^{-4} (2.6×10^{-6})	7.5×10^{-4} (4.8×10^{-6})
			$r = 3$	
.11	6.1×10^{-12}	6.3×10^{-12}	$7.2 \times 10^{-12}(1.0 \times 10^{-14})$	$6.7 \times 10^{-12}(8.7 \times 10^{-15})$
.43	6.8×10^{-10}	8.6×10^{-10}	$7.5 \times 10^{-10}(6.4 \times 10^{-13})$	$7.4 \times 10^{-10}(6.0 \times 10^{-13})$
1.0	5.7×10^{-8}	1.0×10^{-7}	6.3×10^{-8} (1.0×10^{-10})	6.4×10^{-8} (9.6×10^{-10})
2.3	9.4×10^{-6}	2.8×10^{-5}	1.1×10^{-5} (3.2×10^{-8})	1.1×10^{-5} (3.7×10^{-8})
4.0	3.5×10^{-4}	1.6×10^{-3}	4.1×10^{-4} (1.8×10^{-6})	4.5×10^{-4} (2.2×10^{-6})
			$r = 4$	
.11	5.3×10^{-12}	5.5×10^{-12}	$6.2 \times 10^{-12}(8.6 \times 10^{-15})$	$5.9 \times 10^{-12}(7.0 \times 10^{-15})$
.43	6.0×10^{-10}	7.6×10^{-10}	$6.6 \times 10^{-10}(5.4 \times 10^{-13})$	$6.5 \times 10^{-10}(4.8 \times 10^{-13})$
1.0	5.1×10^{-8}	8.9×10^{-8}	5.6×10^{-8} (7.9×10^{-11})	5.6×10^{-8} (6.4×10^{-11})
2.3	8.5×10^{-6}	2.5×10^{-5}	9.6×10^{-6} (2.7×10^{-8})	9.7×10^{-6} (2.3×10^{-8})
4.0	3.2×10^{-4}	1.5×10^{-3}	3.5×10^{-4} (1.1×10^{-6})	3.7×10^{-4} (1.3×10^{-6})
			$r = 5$	
.11	4.7×10^{-12}	4.9×10^{-12}	$5.4 \times 10^{-12}(7.1 \times 10^{-15})$	$5.2 \times 10^{-12}(5.9 \times 10^{-15})$
.43	5.4×10^{-10}	6.8×10^{-10}	$5.9 \times 10^{-10}(4.8 \times 10^{-13})$	$5.8 \times 10^{-10}(3.9 \times 10^{-13})$
1.0	4.6×10^{-8}	8.1×10^{-8}	5.0×10^{-8} (6.7×10^{-11})	4.9×10^{-8} (4.5×10^{-11})
2.3	7.8×10^{-6}	2.3×10^{-5}	8.5×10^{-6} (1.7×10^{-8})	8.5×10^{-6} (1.5×10^{-8})
4.0	2.9×10^{-4}	1.3×10^{-3}	3.2×10^{-4} (7.2×10^{-7})	3.3×10^{-4} (7.7×10^{-7})

Standard Error for the Simulation are in Parentheses

Table 46: Values for the Variance of $\widehat{\boldsymbol{\beta}}$ (4 Mixture Components and 2 Process Variables)

d	Asymptotic	OLS	REML	Pure Error
		$r = 2$		
.11	0.08	0.09	0.09(.00)	0.09(.00)
.43	2.23	3.25	2.65(.00)	2.54(.00)
1.0	35.69	80.65	46.40(.16)	46.60(.14)
2.3	751.67	320.39	1085.94(6.77)	1169.77(69.02)
4.0	6242.75	41467.20	8877.31(70.24)	11478.91(94.98)
		$r = 3$		
.11	0.06	0.07	0.07(.00)	0.07(.00)
.43	1.82	2.67	2.10(.00)	1.98(.00)
1.0	29.77	68.04	36.84(.12)	34.22(.07)
2.3	635.53	2679.31	820.09(45.48)	790.57(33.91)
4.0	5306.77	35863.68	6773.26(47.43)	6938.85(42.87)
		$r = 4$		
.11	0.05	0.05	0.05(.00)	0.05(.00)
.43	1.54	2.25	1.74(.00)	1.64(.00)
1.0	25.53	58.25	30.16(.08)	28.02(.04)
2.3	550.48	2315.20	644.11(27.95)	625.55(18.87)
4.0	4614.86	31104.49	5354.82(27.54)	5270.57(20.95)
		$r = 5$		
.11	0.04	0.05	0.05(.00)	0.05(.00)
.43	1.33	1.94	1.48(.00)	1.40(.00)
1.0	22.35	50.74	25.29(.05)	23.94(.03)
2.3	485.50	2028.05	547.27(19.38)	527.95(10.84)
4.0	4082.56	27307.82	4469.25(16.04)	4467.66(12.82)

Standard Error for the Simulation are in Parentheses

Table 47: Values for the Variance of $\widehat{\boldsymbol{\beta}}$ (4 Mixture Components and 3 Process Variables)

d	Asymptotic	OLS	REML	Pure Error
			$r = 2$	
.11	1.1×10^{-13}	1.2×10^{-13}	$1.2 \times 10^{-13}(5.2 \times 10^{-17})$	$1.1 \times 10^{-13}(4.9 \times 10^{-17})$
.43	2.8×10^{-11}	4.1×10^{-11}	$3.2 \times 10^{-11}(4.7 \times 10^{-14})$	$3.3 \times 10^{-11}(4.4 \times 10^{-13})$
1.0	3.3×10^{-9}	7.5×10^{-9}	4.1×10^{-9} (1.3×10^{-11})	4.6×10^{-9} (1.4×10^{-11})
2.3	6.6×10^{-7}	2.7×10^{-6}	8.2×10^{-7} (4.0×10^{-9})	1.2×10^{-6} (6.7×10^{-9})
4.0	2.6×10^{-5}	1.7×10^{-4}	3.1×10^{-5} (1.7×10^{-7})	5.5×10^{-5} (4.5×10^{-7})
			$r = 3$	
.11	9.2×10^{-14}	9.8×10^{-14}	$9.7 \times 10^{-14}(4.1 \times 10^{-17})$	$9.8 \times 10^{-14}(7.1 \times 10^{-17})$
.43	2.5×10^{-11}	3.6×10^{-11}	$2.7 \times 10^{-11}(3.7 \times 10^{-14})$	$2.7 \times 10^{-11}(2.5 \times 10^{-14})$
1.0	2.9×10^{-9}	6.7×10^{-9}	3.5×10^{-9} (9.8×10^{-12})	3.4×10^{98} (7.5×10^{-12})
2.3	5.9×10^{-7}	2.5×10^{-6}	6.9×10^{-7} (3.0×10^{-9})	7.3×10^{-7} (3.1×10^{-9})
4.0	2.4×10^{-5}	1.6×10^{-4}	2.6×10^{-5} (1.1×10^{-7})	3.1×10^{-5} (1.9×10^{-7})
			$r = 4$	
.11	8.0×10^{-14}	8.5×10^{-14}	$8.4 \times 10^{-14}(3.7 \times 10^{-17})$	$8.8 \times 10^{-14}(8.2 \times 10^{-17})$
.43	2.2×10^{-11}	3.2×10^{-11}	$2.4 \times 10^{-11}(2.9 \times 10^{-14})$	$2.3 \times 10^{-11}(1.5 \times 10^{-14})$
1.0	2.7×10^{-9}	6.1×10^{-9}	3.0×10^{-9} (7.6×10^{-12})	2.9×10^{98} (3.5×10^{-12})
2.3	5.4×10^{-7}	2.3×10^{-6}	5.9×10^{-7} (2.0×10^{-9})	5.8×10^{-7} (1.4×10^{-9})
4.0	2.2×10^{-5}	1.5×10^{-4}	2.3×10^{-5} (7.6×10^{-8})	2.3×10^{-5} (7.6×10^{-8})
			$r = 5$	
.11	7.1×10^{-14}	7.5×10^{-14}	$7.4 \times 10^{-14}(3.3 \times 10^{-17})$	$8.2 \times 10^{-14}(8.6 \times 10^{-17})$
.43	1.9×10^{-11}	2.9×10^{-11}	$2.1 \times 10^{-11}(2.6 \times 10^{-14})$	$2.1 \times 10^{-11}(1.3 \times 10^{-14})$
1.0	2.4×10^{-9}	5.5×10^{-9}	2.7×10^{-9} (5.6×10^{-12})	2.6×10^{-9} (1.8×10^{-12})
2.3	4.9×10^{-7}	2.1×10^{-6}	5.3×10^{-7} (1.3×10^{-9})	5.1×10^{-7} (5.9×10^{-10})
4.0	1.9×10^{-5}	1.3×10^{-4}	2.1×10^{-5} (4.9×10^{-8})	2.1×10^{-5} (3.1×10^{-8})
Standard Error for the Simulation are in Parentheses				

Table 48: Values for the Variance of $\widehat{\beta}$ (3 Mixture Components and 2 Process Variables in a Constrained Region)

d	Asymptotic	OLS	REML	Pure Error
			$r = 2$	
.11	0.38	0.38	0.47(.00)	0.40(.00)
.43	16.27	17.69	18.77(.05)	17.61(.02)
1.0	404.89	502.15	494.88(12.13)	464.89(.61)
2.3	13200.28	21133.80	17230.56(31.84)	17048.69(28.36)
4.0	139796.42	290236.72	195563.56(558.13)	203788.49(536.60)
			$r = 3$	
.11	0.25	0.26	0.27(.00)	0.28(.00)
.43	11.24	12.01	12.30(.03)	12.40(.02)
1.0	286.86	341.64	324.76(.53)	324.30(.60)
2.3	9737.79	14390.76	11437.38(17.68)	11380.29(16.95)
4.0	106541.24	197694.87	130885.49(257.89)	128693.51(235.56)
			$r = 4$	
.11	0.19	0.19	0.20(.00)	0.21(.00)
.43	8.58	9.09	9.14(.01)	9.69(.02)
1.0	222.11	258.87	243.24(.34)	249.69(.45)
2.3	7714.29	10908.96	8691.57(11.12)	8583.57(11.25)
4.0	86067.30	149886.77	99194.52(142.72)	96290.94(127.16)
			$r = 5$	
.11	0.15	0.15	0.16(.00)	0.18(.00)
.43	6.94	7.32	7.28(.01)	8.49(.02)
1.0	181.20	208.38	193.64(.17)	222.79(.58)
2.3	6387.07	8783.52	7032.91(75.21)	7754.71(15.89)
4.0	72193.86	120695.23	81590.13(105.11)	85671.18(135.44)

Standard Error for the Simulation are in Parentheses

OLS performs poorly except when $d < 1$. This is expected because when d is small relative to one, the variance of $\hat{\boldsymbol{\beta}}$ is dominated by the subplot error, σ_ϵ. So, when d is small, the variance-covariance matrix is close to σ_ϵ^2 times an identity matrix which is what OLS is assuming.

The pure error approach and REML are quite comparable when three replicates of the the center of the process variables are run. When only two replicates are run, REML tends to perform better. With four or five replicates, the pure error method performs very well. Therefore, if the experimenter has no prior knowledge of the error variance values and can afford the extra center runs, the pure error method is a simple and effective way to estimate the error variances. To get estimates of the coefficients, once $\mathbf{V}_c$ is calculated, it can be inputted as the weight matrix and weighted least squares can be performed. If REML is to be used, the replicates of the the center of the process variables would most likely not be run. While the design without these points cannot be directly compared to a design with these points, it is believed that REML is helped by the inclusion of the replicate points. An added benefit of including the replicate points and using the pure error approach is the ability to test for lack of fit of the model.

5.2 Second-Order Model in the Process Variables

Suppose now that a second-order model is assumed for the process variables. The combined model using a Taylor series approach is

$$\eta(\mathbf{x}, \mathbf{z}) = \sum_{i=1}^{q} \beta_i^* x_i + \sum\sum_{i<j}^{q} \beta_{ij}^* x_i x_j + \sum\sum_{k<l}^{n} \alpha_{kl} z_k z_l + \sum_{k=1}^{n} \alpha_{kk} z_k^2 + \sum_{i=1}^{q}\sum_{k=1}^{n} \gamma_{ik} x_i z_k \,.$$

Table 49: Relative Efficiencies for Comparing Methods of Estimating $\mathbf{V}$ With 3 Mixture Components and 2 Process Variables

	$r = 2$			$r = 3$		
d	OLS	REML	PE	OLS	REML	PE
.11	1.03	1.06	1.07	1.03	1.06	1.08
.43	1.26	1.13	1.10	1.27	1.12	1.07
1.0	1.78	1.25	1.20	1.79	1.20	1.11
2.3	3.02	1.38	1.40	3.07	1.23	1.18
4.0	4.67	1.48	1.62	4.76	1.33	1.24
	$r = 4$			$r = 5$		
d	OLS	REML	PE	OLS	REML	PE
.11	1.03	1.06	1.08	1.03	1.05	1.08
.43	1.27	1.10	1.06	1.26	1.09	1.05
1.0	1.79	1.13	1.08	1.79	1.13	1.06
2.3	3.06	1.18	1.12	3.05	1.11	1.08
4.0	4.75	1.14	1.13	4.72	1.14	1.08

A design similar to the one presented in the previous section is needed which will support the fit of this model.

Proposed Design

Since a second-order model in the process variables is considered, a ccd is now appropriate. The addition of the axial points will allow for the fitting of this model. Therefore, the design given in Figure 8 will be augmented with axial points. In these axial points, m replicates of the centroid blend will be run. This will preserve a balanced design and provide an additional $2r(m-1)$ df for estimating the subplot error variance in the pure error approach. There are still r replicates of the center of the process variables and m replicates of the centroid in each of these. The design

Table 50: Relative Efficiencies for Comparing Methods of Estimating **V** With 3 Mixture Components and 3 Process Variables

	$r = 2$			$r = 3$		
d	OLS	REML	PE	OLS	REML	PE
.11	1.03	1.21	1.06	1.03	1.18	1.10
.43	1.26	1.22	1.12	1.26	1.10	1.09
1.0	1.79	1.12	1.28	1.77	1.11	1.13
2.3	3.05	1.22	1.60	3.01	1.13	1.21
4.0	4.72	1.24	1.91	4.66	1.16	1.28
	$r = 4$			$r = 5$		
d	OLS	REML	PE	OLS	REML	PE
.11	1.03	1.16	1.10	1.03	1.15	1.09
.43	1.26	1.09	1.07	1.25	1.09	1.06
1.0	1.76	1.09	1.09	1.75	1.08	1.07
2.3	2.98	1.12	1.13	2.95	1.08	1.09
4.0	4.59	1.09	1.16	4.54	1.07	1.10

is shown in Figure 9 for 3 mixture components and 2 process variables. Again, the designs can be extended to higher dimensions with m changing appropriately.

The estimation procedure and simulation study are carried out the same way as discussed in the first-order case. The only addition is that the m replicates of the centroid blend in the axial points contribute to the estimate of the subplot pure error variance.

Comparison of Methods

The same values of d are used and it is still assumed that $\sigma_e^2 = 1$. Also, the designs consisted of $r = 2$, 3, 4 and 5 total replicates of the center of the process variables. The formulas for the ratios of the determinants given for the first-order case still give

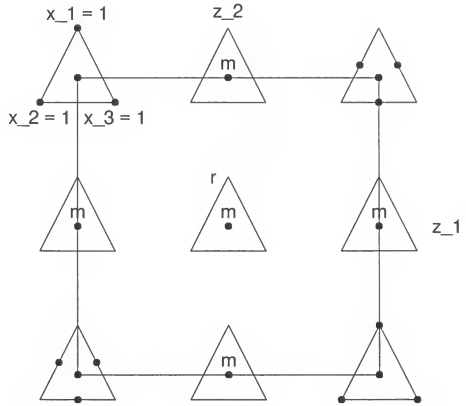


Figure 9: Proposed Design for Split-Plot Structure With a Second-Order Model

Table 51: Relative Efficiencies for Comparing Methods of Estimating **V** With 4 Mixture Components and 2 Process Variables

	$r = 2$			$r = 3$		
d	OLS	REML	PE	OLS	REML	PE
.11	1.06	1.06	1.06	1.06	1.06	1.06
.43	1.45	1.18	1.13	1.46	1.15	1.08
1.0	2.25	1.29	1.30	2.28	1.23	1.14
2.3	4.15	1.44	1.55	4.21	1.29	1.24
4.0	6.64	1.42	1.83	6.75	1.27	1.30
	$r = 4$			$r = 5$		
d	OLS	REML	PE	OLS	REML	PE
.11	1.06	1.06	1.06	1.06	1.05	1.06
.43	1.46	1.12	1.06	1.45	1.11	1.05
1.0	2.28	1.18	1.09	2.26	1.13	1.07
2.3	4.20	1.17	1.13	4.17	1.12	1.08
4.0	6.74	1.16	1.14	6.68	1.09	1.09

the relative efficiencies for the second-order case which are shown in Tables 54 - 58.

OLS is not presented since it performed so poorly in the first-order case and does so in the second-order case as well. REML performs well for all cases, d values, and number of replicates at the center of the process variables. The pure error method does not perform well when there are only 2 replicates at the center of the process variables. With 4 or 5 replicates, the pure error method compares very well with REML.

5.3 Summary

Often times, cost or time constraints make mixture experiments with process variables difficult to run in a completely randomized order. In this situation, fixing

Table 52: Relative Efficiencies for Comparing Methods of Estimating $\mathbf{V}$ With 4 Mixture Components and 3 Process Variables

	$r=2$			$r=3$		
d	OLS	REML	PE	OLS	REML	PE
.11	1.06	1.05	1.04	1.06	1.05	1.06
.43	1.45	1.14	1.16	1.46	1.11	1.08
1.0	2.24	1.23	1.38	2.27	1.16	1.14
2.3	4.11	1.24	1.75	4.19	1.16	1.23
4.0	6.58	1.17	2.12	6.72	1.16	1.28
	$r=4$			$r=5$		
d	OLS	REML	PE	OLS	REML	PE
.11	1.06	1.05	1.10	1.06	1.05	1.15
.43	1.46	1.09	1.06	1.46	1.09	1.07
1.0	2.28	1.13	1.07	2.28	1.10	1.05
2.3	4.21	1.10	1.08	4.20	1.07	1.04
4.0	6.76	1.07	1.07	6.74	1.05	1.03

a level of some factors and then running combinations of the other factors leads to an experiment with a split-plot structure. Designs are proposed which consider the process variables as the whole plot factors and the mixture components as the subplot factors. These designs are an extension to the designs given in Chapter 4.

The split-plot structure of the experiment complicates the estimation of the variance components because OLS is no longer valid. Two alternative methods are presented: REML and a pure error approach. A simulation is conducted to get estimates of the variance components. The two methods along with OLS are compared using the determinant of the variance of $\widehat{\boldsymbol{\beta}}$ and forming a relative efficiency in terms of the asymptotic value. The relative efficiencies give the inflation factor of the size of the confidence ellipsoid around $\boldsymbol{\beta}$, relative to the true size.

Table 53: Relative Efficiencies for Comparing Methods of Estimating
V With 3 Mixture Components and 2 Process Variables in a Constrained Region

	$r = 2$			$r = 3$		
d	OLS	REML	PE	OLS	REML	PE
.11	1.03	1.23	1.07	1.03	1.08	1.08
.43	1.26	1.15	1.10	1.27	1.09	1.07
1.0	1.78	1.22	1.20	1.79	1.13	1.11
2.3	3.02	1.30	1.40	3.07	1.17	1.18
4.0	4.67	1.39	1.62	4.76	1.22	1.24
	$r = 4$			$r = 5$		
d	OLS	REML	PE	OLS	REML	PE
.11	1.03	1.03	1.08	1.03	1.02	1.08
.43	1.27	1.06	1.06	1.26	1.04	1.05
1.0	1.79	1.09	1.08	1.79	1.06	1.06
2.3	3.06	1.12	1.12	3.05	1.10	1.08
4.0	4.75	1.15	1.13	4.72	1.13	1.08

Both first-order and second-order models are considered for the process variables. In the first-order case, the pure error method performs well when there are 3, 4, or 5 replicates of the center of the process variables. REML performs well in all cases and especially when there are only 2 or 3 replicates. With the second-order model, REML performs well across the board while the pure error method performs well only with 4 or 5 replicates. If the experimenter can afford a few extra runs, then the pure error approach is a simple method and has the added feature of being able to test the model for lack of fit.

Table 54: Relative Efficiencies for Comparing Methods of Estimating $\mathbf{V}$ With 3 Mixture Components and 2 Process Variables With a Quadratic Model

	$r = 2$		$r = 3$	
d	REML	PE	REML	PE
.11	1.03	1.04	1.02	1.04
.43	1.07	1.06	1.05	1.04
1.0	1.09	1.14	1.06	1.07
2.3	1.08	1.28	1.06	1.12
4.0	1.09	1.48	1.07	1.15
	$r = 4$		$r = 5$	
d	REML	PE	REML	PE
.11	1.02	1.04	1.02	1.04
.43	1.06	1.03	1.05	1.02
1.0	1.05	1.04	1.05	1.03
2.3	1.05	1.06	1.05	1.04
4.0	1.05	1.07	1.03	1.04

Table 55: Relative Efficiencies for Comparing Methods of Estimating V With 3 Mixture Components and 3 Process Variables With a Quadratic Model

	$r = 2$		$r = 3$	
d	REML	PE	REML	PE
.11	1.25	1.08	1.23	1.07
.43	1.10	1.10	1.09	1.07
1.0	1.06	1.20	1.04	1.11
2.3	1.04	1.39	1.03	1.17
4.0	1.04	1.58	1.03	1.22
	$r = 4$		$r = 5$	
d	REML	PE	REML	PE
.11	1.23	1.06	1.22	1.06
.43	1.09	1.05	1.08	1.04
1.0	1.03	1.07	1.04	1.05
2.3	1.02	1.10	1.02	1.07
4.0	1.02	1.12	1.02	1.07

Table 56: Relative Efficiencies for Comparing Methods of Estimating $\mathbf{V}$ With 4 Mixture Components and 2 Process Variables With a Quadratic Model

	$r = 2$		$r = 3$	
d	REML	PE	REML	PE
.11	1.03	1.04	1.03	1.03
.43	1.07	1.09	1.07	1.05
1.0	1.09	1.20	1.08	1.10
2.3	1.10	1.38	1.06	1.15
4.0	1.08	1.41	1.05	1.17
	$r = 4$		$r = 5$	
d	REML	PE	REML	PE
.11	1.02	1.01	1.03	1.03
.43	1.06	1.03	1.04	1.03
1.0	1.07	1.05	1.04	1.04
2.3	1.04	1.08	1.03	1.04
4.0	1.03	1.08	1.02	1.04

Table 57: Relative Efficiencies for Comparing Methods of Estimating V With 4 Mixture Components and 3 Process Variables With a Quadratic Model

	$r = 2$		$r = 3$	
d	REML	PE	REML	PE
.11	1.03	1.04	1.02	1.03
.43	1.05	1.09	1.05	1.05
1.0	1.05	1.20	1.04	1.09
2.3	1.04	1.39	1.03	1.13
4.0	1.02	1.52	1.02	1.18
	$r = 4$		$r = 5$	
d	REML	PE	REML	PE
.11	1.02	1.03	1.03	1.03
.43	1.04	1.03	1.04	1.04
1.0	1.03	1.05	1.03	1.04
2.3	1.02	1.07	1.02	1.07
4.0	1.01	1.07	1.01	1.06

Table 58: Relative Efficiencies for Comparing Methods of Estimating $\mathbf{V}$ With 3 Mixture Components and 2 Process Variables in a Constrained Region With a Quadratic Model

	$r = 2$		$r = 3$	
d	REML	PE	REML	PE
.11	1.06	1.14	1.06	1.10
.43	1.11	1.16	1.08	1.11
1.0	1.14	1.23	1.12	1.14
2.3	1.21	1.40	1.16	1.22
4.0	1.22	1.61	1.17	1.30
	$r = 4$		$r = 5$	
d	REML	PE	REML	PE
.11	1.07	1.09	1.05	1.06
.43	1.07	1.09	1.06	1.06
1.0	1.10	1.11	1.08	1.08
2.3	1.12	1.15	1.09	1.12
4.0	1.13	1.19	1.10	1.15

CHAPTER 6
SUMMARY AND CONCLUSIONS

Many experiments that are performed in industry have cost and/or time constraints which limit the size of the experiment. A concern with small experiments is how much information can really be obtained in this setting. The global theme of this dissertation is to propose methods for constructing designs for small experiments that will give as much information as possible about the factors involved.

In Chapter 3, the experiments of interest are those that contain some factors that are hard-to-change and some that are relatively easy-to-change. When this is the case, it is more economical to run the experiment as a split-plot experiment. If all of the factors are at two levels, the size of the experiment can be reduced by using fractional factorials, split-plot confounding, or both. We have presented methods for constructing sixteen run designs that will be economically feasible. Also, if the experiment is in the framework of robust parameter design, then these designs handle the subplot factors quite well. However, if possible, it is beneficial from an estimation point of view to convince the experimenter to add eight more runs at the subplot level. This will break alias chains in order to get better estimates of the parameters as well as add some degrees of freedom for testing. Chapter 3 is concluded with several examples of the proposed methods.

Mixture experiments are also widely used in industry. Often times the quality of the product depends not only on the relative proportions of the mixture components

but also on the processing conditions. We considered the process variables as two-level factors. A new second order model that is intended to be a compromise between the common mixture models used and standard response surface models is given in Chapter 4. With this model in hand, a new class of designs is proposed that will support the fit of this model. Keeping to the global theme, the designs are kept relatively small. Assuming a completely randomized design, the analysis for these experiments is provided with a small discussion of lack of fit.

Depending on the types of process variables being used, it may be too costly to run a completely randomized experiment. In Chapter 5, we consider the experiment as a split-plot. The process variables are assigned as the whole plot factors with the mixture components serving as the subplot factors. The split-plot nature complicates these experiments since ordinary least squares is no longer valid. An iterative approach is needed. We investigate how ordinary least squares holds up to restricted maximum likelihood under various ratios of the whole plot error variance to the subplot error variance. Also, another method for estimating these variance components is presented which uses pure error from replicated points.

In consideration of the types of experiments that are being run in industry, the size of the experiment has been of great concern in this dissertation. We have presented methods for obtaining as much information as possible with a small experiment. Many industrial experiments fall into the class of split-plot experiments either by nature or by consequence of cost. We have shown how to design and analyze small experiments which are in this class.

APPENDIX A
TABLES FOR CHAPTER 3 DESIGNS

The tables in this Appendix give the design points for the seven cases discussed in Chapter 3. The coding convention is as follows:

- -1 is the low level of a factor
- 1 is the high level of a factor.

The points for 16 run designs are listed at the beginning of the table followed by the additional 8 points.

Table 59: Design Points for 2 WP Factors and 4 SP Factors Using Separate Fractions

A	B	P	Q	R	S
1	−1	−1	1	−1	−1
1	−1	−1	−1	1	−1
1	−1	1	−1	−1	1
1	−1	1	1	1	1
−1	1	−1	1	−1	−1
−1	1	−1	−1	1	−1
−1	1	1	−1	−1	1
−1	1	1	1	1	1
1	1	−1	1	−1	−1
1	1	−1	−1	1	−1
1	1	1	−1	−1	1
1	1	1	1	1	1
−1	−1	−1	1	−1	−1
−1	−1	−1	−1	1	−1
−1	−1	1	−1	−1	1
−1	−1	1	1	1	1
1	−1	−1	−1	−1	1
1	−1	−1	1	1	1
−1	1	1	1	−1	−1
−1	1	1	−1	1	−1
1	1	1	−1	−1	−1
1	1	1	1	1	−1
−1	−1	−1	1	−1	1
−1	−1	−1	−1	1	1

Table 60: Design Points for 2 WP Factors and 4 SP Factors Using Split-Plot Confounding

A	B	P	Q	R	S
1	−1	1	−1	−1	−1
1	−1	1	1	1	−1
1	−1	−1	1	−1	1
1	−1	−1	−1	1	1
−1	1	−1	−1	−1	1
−1	1	1	1	−1	−1
−1	1	1	−1	1	−1
−1	1	−1	1	1	1
1	1	−1	1	−1	−1
1	1	−1	−1	1	−1
1	1	1	−1	−1	1
1	1	1	1	1	1
−1	−1	−1	1	1	−1
−1	−1	1	1	−1	1
−1	−1	1	−1	1	1
−1	−1	−1	−1	−1	−1
1	−1	−1	−1	−1	−1
1	−1	−1	1	1	−1
−1	1	1	−1	−1	1
−1	1	1	1	1	1
1	1	1	−1	−1	−1
1	1	1	1	1	−1
−1	−1	−1	1	1	1
−1	−1	−1	−1	−1	1

Table 61: Design Points for 3 WP Factors and 3 SP Factors Using Separate Fractions

A	B	C	P	Q	R
1	−1	−1	1	−1	−1
1	−1	−1	−1	1	−1
1	−1	−1	−1	−1	1
1	−1	−1	1	1	1
−1	1	−1	1	−1	−1
−1	1	−1	−1	1	−1
−1	1	−1	−1	−1	1
−1	1	−1	1	1	1
−1	−1	1	1	−1	−1
−1	−1	1	−1	1	−1
−1	−1	1	−1	−1	1
−1	−1	1	1	1	1
1	1	1	1	−1	−1
1	1	1	−1	1	−1
1	1	1	−1	−1	1
1	1	1	1	1	1
1	−1	−1	−1	−1	−1
1	−1	−1	−1	1	1
−1	1	−1	1	1	−1
−1	1	−1	1	−1	1
−1	−1	1	1	1	−1
−1	−1	1	1	−1	1
1	1	1	−1	−1	−1
1	1	1	−1	1	1

Table 62: Design Points for 3 WP Factors and 3 SP Factors Using Split-Plot Confounding

A	B	C	P	Q	R
1	−1	−1	1	1	−1
1	−1	−1	1	−1	1
1	−1	−1	−1	1	1
1	−1	−1	−1	−1	−1
−1	1	−1	1	1	−1
−1	1	−1	1	−1	1
−1	1	−1	−1	1	1
−1	1	−1	−1	−1	−1
−1	−1	1	1	−1	−1
−1	−1	1	−1	1	−1
−1	−1	1	−1	−1	1
−1	−1	1	1	1	1
1	1	1	1	−1	−1
1	1	1	−1	1	−1
1	1	1	−1	−1	1
1	1	1	1	1	1
1	−1	−1	−1	1	−1
1	−1	−1	−1	−1	1
−1	1	−1	1	1	1
−1	1	−1	1	−1	−1
−1	−1	1	1	1	−1
−1	−1	1	1	−1	1
1	1	1	−1	−1	−1
1	1	1	−1	1	1

Table 63: Design Points for 3 WP Factors and 4 SP Factors Using Separate Fractions

A	B	C	P	Q	R	S
1	−1	−1	−1	1	−1	−1
1	−1	−1	−1	−1	1	−1
1	−1	−1	1	−1	−1	1
1	−1	−1	1	1	1	1
−1	1	−1	−1	1	−1	−1
−1	1	−1	−1	−1	1	−1
−1	1	−1	1	−1	−1	1
−1	1	−1	1	1	1	1
−1	−1	1	−1	1	−1	−1
−1	−1	1	−1	−1	1	−1
−1	−1	1	1	−1	−1	1
−1	−1	1	1	1	1	1
1	1	1	−1	1	−1	−1
1	1	1	−1	−1	1	−1
1	1	1	1	−1	−1	1
1	1	1	1	1	1	1
1	−1	−1	−1	−1	−1	1
1	−1	−1	−1	1	1	1
−1	1	−1	1	1	−1	−1
−1	1	−1	1	−1	1	−1
−1	−1	1	1	−1	−1	−1
−1	−1	1	1	1	1	−1
1	1	1	−1	1	−1	1
1	1	1	−1	−1	1	1

Table 64: Design Points for 3 WP Factors and 4 SP Factors Using Split-Plot Confounding

A	B	C	P	Q	R	S
1	−1	−1	1	−1	−1	−1
1	−1	−1	1	1	1	−1
1	−1	−1	−1	1	−1	1
1	−1	−1	−1	−1	1	1
−1	1	−1	−1	−1	−1	1
−1	1	−1	1	1	−1	−1
−1	1	−1	1	−1	1	−1
−1	1	−1	−1	1	1	1
−1	−1	1	−1	1	1	−1
−1	−1	1	1	1	−1	1
−1	−1	1	1	−1	1	1
−1	−1	1	−1	−1	−1	−1
1	1	1	−1	1	−1	−1
1	1	1	−1	−1	1	−1
1	1	1	1	−1	−1	1
1	1	1	1	1	1	1
1	−1	−1	−1	−1	−1	−1
1	−1	−1	−1	1	1	−1
−1	1	−1	1	−1	−1	1
−1	1	−1	1	1	1	1
−1	−1	1	−1	1	1	1
−1	−1	1	−1	−1	−1	1
1	1	1	1	−1	−1	−1
1	1	1	1	1	1	−1

Table 65: Design Points for 4 WP Factors and 2 SP Factors Using Separate Fractions

A	B	C	D	P	Q
−1	1	−1	−1	1	−1
−1	1	−1	−1	−1	1
−1	1	−1	−1	1	1
−1	1	−1	−1	−1	−1
−1	−1	1	−1	1	−1
−1	−1	1	−1	−1	1
−1	−1	1	−1	1	1
−1	−1	1	−1	−1	−1
1	−1	−1	1	1	−1
1	−1	−1	1	−1	1
1	−1	−1	1	1	1
1	−1	−1	1	−1	−1
1	1	1	1	1	−1
1	1	1	1	−1	1
1	1	1	1	1	1
1	1	1	1	−1	−1
−1	−1	−1	1	1	−1
−1	−1	−1	1	−1	1
−1	−1	−1	1	1	1
−1	−1	−1	1	−1	−1
−1	1	1	1	1	−1
−1	1	1	1	−1	1
−1	1	1	1	1	1
−1	1	1	1	−1	−1

Table 66: Design Points for 4 WP Factors and 3 SP Factors Using Separate Fractions

A	B	C	D	P	Q	R
−1	1	−1	−1	1	−1	−1
−1	1	−1	−1	−1	1	−1
−1	1	−1	−1	−1	−1	1
−1	1	−1	−1	1	1	1
−1	−1	1	−1	1	−1	−1
−1	−1	1	−1	−1	1	−1
−1	−1	1	−1	−1	−1	1
−1	−1	1	−1	1	1	1
1	−1	−1	1	1	−1	−1
1	−1	−1	1	−1	1	−1
1	−1	−1	1	−1	−1	1
1	−1	−1	1	1	1	1
1	1	1	1	1	−1	−1
1	1	1	1	−1	1	−1
1	1	1	1	−1	−1	1
1	1	1	1	1	1	1
−1	−1	−1	1	1	−1	−1
−1	−1	−1	1	−1	1	−1
−1	−1	−1	1	−1	−1	1
−1	−1	−1	1	1	1	1
−1	1	1	1	1	1	−1
−1	1	1	1	1	−1	1
−1	1	1	1	−1	1	1
−1	1	1	1	−1	−1	−1

Table 67: Design Points for 4 WP Factors and 3 SP Factors Using Split-Plot Confounding

A	B	C	D	P	Q	R
−1	1	−1	−1	1	1	−1
−1	1	−1	−1	1	−1	1
−1	1	−1	−1	−1	1	1
−1	1	−1	−1	−1	−1	−1
−1	−1	1	−1	1	−1	−1
−1	−1	1	−1	−1	1	−1
−1	−1	1	−1	−1	−1	1
−1	−1	1	−1	1	1	1
1	−1	−1	1	1	1	−1
1	−1	−1	1	1	−1	1
1	−1	−1	1	−1	1	1
1	−1	−1	1	−1	−1	−1
1	1	1	1	1	−1	−1
1	1	1	1	−1	1	−1
1	1	1	1	−1	−1	1
1	1	1	1	1	1	1
−1	−1	−1	1	1	−1	−1
−1	−1	−1	1	−1	1	−1
−1	−1	−1	1	−1	−1	1
−1	−1	−1	1	1	1	1
−1	1	1	1	1	1	−1
−1	1	1	1	1	−1	1
−1	1	1	1	−1	1	1
−1	1	1	1	−1	−1	−1

Table 68: Design Points for 4 WP Factors and 4 SP Factors Using Same Fraction

A	B	C	D	P	Q	R	S
−1	1	−1	−1	−1	1	−1	−1
−1	1	−1	−1	−1	−1	1	−1
−1	1	−1	−1	1	−1	−1	1
−1	1	−1	−1	1	1	1	1
−1	−1	1	−1	−1	1	−1	−1
−1	−1	1	−1	−1	−1	1	−1
−1	−1	1	−1	1	−1	−1	1
−1	−1	1	−1	1	1	1	1
1	−1	−1	1	−1	1	−1	−1
1	−1	−1	1	−1	−1	1	−1
1	−1	−1	1	1	−1	−1	1
1	−1	−1	1	1	1	1	1
1	1	1	1	−1	1	−1	−1
1	1	1	1	−1	−1	1	−1
1	1	1	1	1	−1	−1	1
1	1	1	1	1	1	1	1
−1	−1	−1	1	1	1	−1	−1
−1	−1	−1	1	1	−1	1	−1
−1	−1	−1	1	−1	−1	−1	1
−1	−1	−1	1	−1	1	1	1
−1	1	1	1	−1	1	−1	1
−1	1	1	1	−1	−1	1	1
−1	1	1	1	1	−1	−1	−1
−1	1	1	1	1	1	1	−1

Table 69: Design Points for 4 WP Factors and 4 SP Factors Using Split-Plot Confounding

A	B	C	D	P	Q	R	S
−1	1	−1	−1	1	−1	−1	−1
−1	1	−1	−1	1	1	1	−1
−1	1	−1	−1	−1	1	−1	1
−1	1	−1	−1	−1	−1	1	1
−1	−1	1	−1	−1	1	1	1
−1	−1	1	−1	−1	−1	−1	1
−1	−1	1	−1	1	1	−1	−1
−1	−1	1	−1	1	−1	1	−1
1	−1	−1	1	−1	−1	−1	−1
1	−1	−1	1	−1	1	1	−1
1	−1	−1	1	1	1	−1	1
1	−1	−1	1	1	−1	1	1
1	1	1	1	−1	1	−1	−1
1	1	1	1	−1	−1	1	−1
1	1	1	1	1	−1	−1	1
1	1	1	1	1	1	1	1
−1	−1	−1	1	1	1	1	1
−1	−1	−1	1	1	−1	−1	1
−1	−1	−1	1	−1	1	−1	−1
−1	−1	−1	1	−1	−1	1	−1
−1	1	1	1	−1	1	−1	1
−1	1	1	1	−1	−1	1	1
−1	1	1	1	1	−1	−1	−1
−1	1	1	1	1	1	1	−1

APPENDIX B
TABLES FOR CHAPTER 4 DESIGNS

The tables in this Appendix give the design points for the two proposed designs and all cases discussed in Chapter 4. The coding convention is as follows:

- the mixture components are on a scale of 0 to 1 and the decimal values refer to proportions or fractions of the components
- -1 is the low level of a process variable
- 0 is the middle level of a process variable
- 1 is the high level of a process variable.

The additional design points necessary for the simplex-centroid at the center of the process variable are given below the design using just the centroid at the center.

Table 70: Design Points for 3-2 Case

x_1	x_2	x_3	z_1	z_2
1	0	0	−1	1
1	0	0	1	−1
0	1	0	−1	1
0	1	0	1	−1
0	0	1	−1	1
0	0	1	1	−1
.5	.5	0	−1	−1
.5	.5	0	1	1
.5	0	.5	−1	−1
.5	0	.5	1	1
0	.5	.5	−1	−1
0	.5	.5	1	1
.33	.33	.33	−1	0
.33	.33	.33	1	0
.33	.33	.33	0	−1
.33	.33	.33	0	1
.33	.33	.33	0	0
Additional Points for Simplex-Centroid				
1	0	0	0	0
0	1	0	0	0
0	0	1	0	0
.5	.5	0	0	0
.5	0	.5	0	0
0	.5	.5	0	0

Table 71: Design Points for 3-3 Case

x_1	x_2	x_3	z_1	z_2	z_3
1	0	0	−1	−1	1
1	0	0	−1	1	−1
1	0	0	1	1	1
1	0	0	1	−1	−1
0	1	0	−1	−1	1
0	1	0	−1	1	−1
0	1	0	1	1	1
0	1	0	1	−1	−1
0	0	1	−1	−1	1
0	0	1	−1	1	−1
0	0	1	1	1	1
0	0	1	1	−1	−1
.5	.5	0	−1	−1	−1
.5	.5	0	−1	1	1
.5	.5	0	1	−1	1
.5	.5	0	1	1	−1
.5	0	.5	−1	−1	−1
.5	0	.5	−1	1	1
.5	0	.5	1	−1	1
.5	0	.5	1	1	−1
0	.5	.5	−1	−1	−1
0	.5	.5	−1	1	1
0	.5	.5	1	−1	1
0	.5	.5	1	1	−1
.33	.33	.33	−1	0	0
.33	.33	.33	1	0	0
.33	.33	.33	0	−1	0
.33	.33	.33	0	1	0
.33	.33	.33	0	0	−1
.33	.33	.33	0	0	1
.33	.33	.33	0	0	0
Additional Points for Simplex-Centroid					
1	0	0	0	0	0
0	1	0	0	0	0
0	0	1	0	0	0
.5	.5	0	0	0	0
.5	0	.5	0	0	0
0	.5	.5	0	0	0

Table 72: Design Points for 4-2 Case

x_1	x_2	x_3	x_4	z_1	z_2
1	0	0	0	−1	1
1	0	0	0	1	−1
0	1	0	0	−1	1
0	1	0	0	1	−1
0	0	1	0	−1	1
0	0	1	0	1	−1
0	0	0	1	−1	1
0	0	0	1	1	−1
.5	.5	0	0	−1	−1
.5	.5	0	0	1	1
.5	0	.5	0	−1	−1
.5	0	.5	0	1	1
.5	0	0	.5	−1	−1
.5	0	0	.5	1	1
0	.5	.5	0	−1	−1
0	.5	.5	0	1	1
0	.5	0	.5	−1	−1
0	.5	0	.5	1	1
0	0	.5	.5	−1	−1
0	0	.5	.5	1	1
.25	.25	.25	.25	−1	0
.25	.25	.25	.25	1	0
.25	.25	.25	.25	0	−1
.25	.25	.25	.25	0	1
.25	.25	.25	.25	0	0
Additional Points for Simplex-Centroid					
1	0	0	0	0	0
0	1	0	0	0	0
0	0	1	0	0	0
0	0	0	1	0	0
.5	.5	0	0	0	0
.5	0	.5	0	0	0
.5	0	0	.5	0	0
0	.5	.5	0	0	0
0	.5	0	.5	0	0
0	0	.5	.5	0	0

Table 73: Design Points for 4-3 Case

x_1	x_2	x_3	x_4	z_1	z_2	z_3	x_1	x_2	x_3	x_4	z_1	z_2	z_3
1	0	0	0	−1	−1	1	0	0	.5	.5	−1	−1	−1
1	0	0	0	−1	1	−1	0	0	.5	.5	−1	1	1
1	0	0	0	1	1	1	0	0	.5	.5	1	−1	1
1	0	0	0	1	−1	−1	0	0	.5	.5	1	1	−1
0	1	0	0	−1	−1	1	.25	.25	.25	.25	−1	0	0
0	1	0	0	−1	1	−1	.25	.25	.25	.25	1	0	0
0	1	0	0	−1	1	−1	.25	.25	.25	.25	0	−1	0
0	1	0	0	1	1	1	.25	.25	.25	.25	0	1	0
0	0	1	0	−1	−1	1	.25	.25	.25	.25	0	0	−1
0	0	1	0	−1	1	−1	.25	.25	.25	.25	0	0	1
0	0	1	0	1	1	1	.25	.25	.25	.25	0	0	0
0	0	1	0	1	−1	−1	Additional Points for						
0	0	0	1	−1	−1	1	Simplex-Centroid						
0	0	0	1	−1	1	−1	1	0	0	0	0	0	0
0	0	0	1	1	1	1	0	1	0	0	0	0	0
0	0	0	1	1	−1	−1	0	0	1	0	0	0	0
.5	.5	0	0	−1	−1	−1	0	0	0	1	0	0	0
.5	.5	0	0	−1	1	1	.5	.5	0	0	0	0	0
.5	.5	0	0	1	−1	1	.5	0	.5	0	0	0	0
.5	.5	0	0	1	1	−1	.5	0	0	.5	0	0	0
.5	0	.5	0	−1	−1	−1	0	.5	.5	0	0	0	0
.5	0	.5	0	−1	1	1	0	.5	0	.5	0	0	0
.5	0	.5	0	1	−1	1	0	0	.5	.5	0	0	0
.5	0	.5	0	1	1	−1							
.5	0	0	.5	−1	−1	−1							
.5	0	0	.5	−1	1	1							
.5	0	0	.5	1	−1	1							
.5	0	0	.5	1	1	−1							
0	.5	.5	0	−1	−1	−1							
0	.5	.5	0	−1	1	1							
0	.5	.5	0	1	−1	1							
0	.5	.5	0	1	1	−1							
0	.5	0	.5	−1	−1	−1							
0	.5	0	.5	−1	1	1							
0	.5	0	.5	1	−1	1							
0	.5	0	.5	1	1	−1							

Table 74: Design Points for 3-2 Case with Upper and Lower Constraints

x_1'	x_2'	x_3'	z_1	z_2	x_1'	x_2'	x_3'	z_1	z_2
0	.4	.6	1	−1	Additional Points for				
0	.4	.6	−1	1	Simplex-Centroid				
0	.6	.4	−1	1	0	.4	.6	0	0
0	.6	.4	1	−1	0	.6	.4	0	0
.4	.6	0	1	−1	.4	.6	0	0	0
.4	.6	0	−1	1	.4	0	.6	0	0
.4	0	.6	1	−1	.6	.4	0	0	0
.4	0	.6	−1	1	.6	0	.4	0	0
.6	.4	0	−1	1	.5	.5	0	0	0
.6	.4	0	1	−1	.5	0	.5	0	0
.6	0	.4	−1	1	0	.5	.5	0	0
.6	0	.4	1	−1	.2	.2	.6	0	0
0	.5	.5	−1	−1	.2	.6	.2	0	0
0	.5	.5	1	1	.6	.2	.2	0	0
.6	.2	.2	−1	−1					
.6	.2	.2	1	1					
.5	0	.5	−1	−1					
.5	0	.5	1	1					
.2	.6	.2	−1	−1					
.2	.6	.2	1	1					
.5	.5	0	−1	−1					
.5	.5	0	1	1					
.2	.2	.6	−1	−1					
.2	.2	.6	1	1					
.33	.33	.33	−1	0					
.33	.33	.33	1	0					
.33	.33	.33	0	−1					
.33	.33	.33	0	1					
.33	.33	.33	0	0					

APPENDIX C
SAS CODE FOR PROC MIXED

Consider the model of the form

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\delta} + \boldsymbol{\epsilon}$$

where $\boldsymbol{\delta} + \boldsymbol{\epsilon} \sim N(0, \mathbf{V})$ and $\mathbf{V} = \sigma_\delta^2\mathbf{J} + \sigma_\epsilon^2\mathbf{I}$. Then, the following code can be used to obtain estimates of σ_δ^2 and σ_ϵ^2.

PROC MIXED METHOD = REML;

CLASS WP;

MODEL Y = FIXED EFFECTS;

RANDOM WP;

where WP is a classification variable defining into which whole plot each observation falls, the model statement defines the model matrix, $\mathbf{X}$.

REFERENCES

[1] ADDELMAN, S. (1964). "Some Two-Level Factorial Plans With Split-Plot Confounding." *Technometrics* 30, pp. 253–258.

[2] ANDERSON, R. L. (1952). *Statistical Theory in Research.* McGraw-Hill, New York, NY.

[3] BARTLETT, M. S. (1935). Discussion of "Complex Experiments" by F. Yates. *Journal of the Royal Statistical Society*, Ser. B, 2, pp. 224–226.

[4] BINGHAM, D. and SITTER, R. S. (1999). "Minimum-Aberration Two-Level Fractional Factorial Split-Plot Designs." *Technometrics* 41, pp. 62–70.

[5] BISGAARD, S. (1999). "The Design and Analysis of $2^{k-p} \times 2^{q-r}$ Split Plot Experiments." accepted by *Journal of Quality Technology.*

[6] BOX. G. E. P. and DRAPER. N. R. (1987). *Empirical Model Building and Response Surfaces.* John Wiley & Sons, New York, NY.

[7] BOX, G. E. P. and HUNTER, J. S. (1961a). "The 2^{k-p} Fractional Factorial Designs, I." *Technometrics* 3. pp. 311–352.

[8] BOX, G. E. P. and HUNTER. J. S. (1961b). "The 2^{k-p} Fractional Factorial Designs, II." *Technometrics* 3. pp. 449–458.

[9] BOX, G. E. P. and JONES, S. (1992). "Split-Plot Designs for Robust Product Experimentation." *Journal of Applied Statistics* 19, pp. 3–26.

[10] BOX, G. E, P. and WILSON, K. B. (1951). "On the Experimental Attainment of Optimum Conditions." *Journal of the Royal Statistical Society.* Ser. B, 13, pp. 1–45.

[11] CANTELL, B. and RAMIREZ, J. G. (1994). "Robust Design of a Polysilicon Deposition Process Using a Split-Plot Analysis." *Quality Reliability Engineering International* 10, pp. 123–132.

[12] CORNELL, J. A. (1990). *Experiments With Mixtures: Designs, Models, and the Analysis of Mixture Data* 2nd ed. John Wiley & Sons, New York. NY.

[13] CORNELL, J. A. (1988). "Analyzing Data from Mixture Experiments Containing Process Variables: A Split-Plot Approach. *Journal of Quality Technology* 20, pp. 2–23.

[14] CORNELL, J. A. and GORMAN, J. W. (1984). "Fractional Design Plans for Process Variables in Mixture Experiments." *Journal of Quality Technology* 16, pp. 20–38.

[15] COX, D. R. (1958). *Planning Experiments*. John Wiley & Sons. New York, NY.

[16] CZITROM, V. (1988). "Mixture Experiments with Process Variables: D-Optimal Orthogonal Experimental Designs." *Communications in Statistics—Theory and Methods* 17, pp. 105–121.

[17] CZITROM, V. (1989). "Experimental Designs for Four Mixture Components with Process Variables." *Communications in Statistics—Theory and Methods* 18, pp. 4561–4581.

[18] DRAPER, N. R. ; PRESCOTT, P. ; LEWIS, S. M. ; DEAN, A. M. ; JOHN, P. W. M. ; and TUCK, M. G. (1993). "Mixture Designs for Four Components in Orthogonal Blocks." *Technometrics* 35, pp. 268–276.

[19] FINNEY, D. J. (1945). "The Fractional Replication of Factorial Arrangements." *Annals of Eugenics* 12, pp. 291–301.

[20] FISHER, R. A. (1926). "The Arrangement of Field Experiments." *Journal of the Ministry of Agriculture* 33. pp. 503–513.

[21] HUANG, P. ; CHEN, D. and VOELKEL, J. O. (1998). "Minimum-Aberration Two-Level Split-Plot Designs." *Technometrics* 40, pp. 314–326.

[22] KEMPTHORNE, O. (1952). *The Design and Analysis of Experiments*. John Wiley & Sons, New York, NY.

[23] KHURI, A. I. and CORNELL, J. A. (1996). *Response Surfaces* 2nd edition. Marcel Dekker, New York. NY.

[24] KIEFER, J. and WOLFOWITZ, J. (1959). "Optimum Designs in Regression Problems." *Annals of Mathematical Statistics* 30, pp. 271–294.

[25] LETSINGER, J. D. ; MYERS, R. H. ; and LENTNER, M. (1996). "Response Surface Methods for Bi-Randomization Structures." *Journal of Quality Technology* 28, pp. 381-397.

[26] LUCAS, J. M. and JU, H. L. (1992). "Split Plotting and Randomization in Industrial Experiments." *ASQC Quality Congress Transactions*. American Society for Quality Control, Nashville. TN. pp. 374–382.

[27] MARTIN, F. G. (1992). *Statistical Design and Analysis.*

[28] MEE, R. W. and BATES, R. L. (1998). "Split-Lot Designs: Experiments for Multistage Batch Processes." *Technometrics* 40. pp. 127–140.

[29] MILLER, A. (1997). "Strip-Plot Configurations of Fractional Factorials." *Technometrics* 39. pp. 153–161.

[30] MITCHELL, T. J. (1974). "An Algorithm for the Construction of D-Optimal Experimental Designs." *Technometrics* 20, pp. 203–210.

[31] MONTGOMERY, D. C. (1997). *Design and Analysis of Experiments* 4th edition. John Wiley & Sons, New York, NY.

[32] MYERS, R. H. and MONTGOMERY, D. C. (1995). *Response Surface Methodology: Process and Product Optimization Using Designed Experiments.* John Wiley & Sons, New York, NY.

[33] O'DONNELL, E. M. and VINING, G. G. (1997). "Mean Squared Error of Prediction Approach to the Analysis of the Combined Array." *Journal of Applied Statistics* 24, pp. 733–746.

[34] PHADKE, M. S. (1989). *Qualtiy Engineering Using Robust Design.* Prentice Hall, Englewood Cliffs, NJ.

[35] PLACKETT, R. L. and BURMAN, J. T. (1946). "The Design of Optimum Multifactorial Experiments." *Biometrika* 33, pp. 305–325.

[36] ROBINSON, J. (1967). "Incomplete Split Plot Designs." *Biometrics* 23. pp. 793–802.

[37] ROBINSON, J. (1970). "Blocking in Incomplete Split Plot Designs." *Biometrika* 57, pp. 347–350.

[38] RUSSELL, T. S. and BRADLEY, R. A. (1958). "One-Way Variances in the Two-Way Classification." *Biometrika* 45, pp. 111–129.

[39] SANTER, T. J. and PAN, G. (1997). "Subset Selection in Two-Factor Experiments Using Randomization Restricted Designs." *Journal of Statistical Planning and Inference* 62, pp. 339–363.

[40] SAS INSTITUTE (1989). *SAS/QC Software: Reference Guide, Ver.6.* 1st ed. SAS Institute, Cary, NC.

[41] SHOEMAKER, A. C. ; TSUI, K. ; and WU, C. F. J. (1991). "Economical Experimentation Methods for Robust Designs." *Technometrics* 33. pp. 415–427.

[42] SNEE, R. D. (1985). "Computer-Aided Design of Experiments—Some Practical Experience." *Journal of Quality Technology* 17, pp. 222–236.

[43] TAGUCHI, G. (9187). *Systems of Experimental Design, Vol. 1 and 2.* Kraus International Publications, White Plains, NY.

[44] WELCH, W. J. ; KANG, T. ; and SACKS, J. (1990). "Computer Experiments for Quality Control by Parameter Design." *Journal of Quality Technology* 22, pp. 15–22.

[45] WOODING, W. M. (1973). "The Split-Plot Design." *Journal of Quality Technology* 5, pp. 16–33.

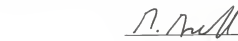
[46] YATES, F. (1937). *The Design and Analysis of Factorial Experiments.* Imperial Bureau of Soil Sciences, Harpenden, England.

BIOGRAPHICAL SKETCH

Scott Kowalski was born on February 5, 1969, in Kitchener, Ontario, Canada. He received his bachelor's degree in statistics at the University of Florida in 1993 and his master's degree in staistics also from the University of Florida. Under the supervision of Professor G. Geoffrey Vining, he received a Doctor of Philosophy degree in statistics in December 1999.

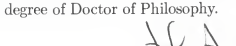
He has been married to Kim since 1994. They have a beautiful daughter, Regan Michelle.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



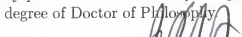
G. Geoffrey Vining, Chairman
Associate Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



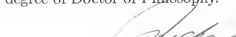
John A. Cornell
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



James P. Hobert
Assistant Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Richard L. Scheaffer
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Diane A. Schaub
Assistant Professor of Industrial
and Systems Engineering

This dissertation was submitted to the Graduate Faculty of the Department of Statistics in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 1999

Dean, Graduate School